

Non-local gravity and comparison with observational datasets.

II. Updated results and Bayesian model comparison with Λ CDM¹

Yves Dirian^a, Stefano Foffa^a, Martin Kunz^{a,b}, Michele Maggiore^a and Valeria Pettorino^c

^a*Département de Physique Théorique and Center for Astroparticle Physics,
Université de Genève, 24 quai Ansermet, CH-1211 Genève 4, Switzerland*

^b*African Institute for Mathematical Sciences, 6 Melrose Road, Muizenberg, 7945, South Africa*

^c*Institut für Theoretische Physik, Universität Heidelberg, Philosophenweg 16, D-69120 Heidelberg*

Abstract

We present a comprehensive and updated comparison with cosmological observations of two non-local modifications of gravity previously introduced by our group, the so called RR and RT models. We implement the background evolution and the cosmological perturbations of the models in a modified Boltzmann code, using CLASS. We then test the non-local models against the *Planck* 2015 TT, TE, EE and Cosmic Microwave Background (CMB) lensing data, isotropic and anisotropic Baryonic Acoustic Oscillations (BAO) data, JLA supernovae, H_0 measurements and growth rate data, and we perform Bayesian parameter estimation. We then compare the RR, RT and Λ CDM models, using the Savage-Dickey method. We find that the RT model and Λ CDM perform equally well, while the performance of the RR model with respect to Λ CDM depends on whether or not we include a prior on H_0 based on local measurements.

¹Based on observations obtained with Planck (<http://www.esa.int/Planck>), an ESA science mission with instruments and contributions directly funded by ESA Member States, NASA, and Canada.

1 Introduction

In the last two decades Λ CDM has gradually become the standard cosmological paradigm. Indeed, Λ CDM fits all cosmological observations with great precision, with a limited set of free parameters. Still, the presence of the cosmological constant raises a number of conceptual issues such as the coincidence problem, and the related fact that a cosmological constant is not technically natural from the point of view of the stability under radiative corrections. Thus, much effort is being devoted to looking for alternatives, see e.g. [1] for a recent review. Most importantly, the wealth of high-quality cosmological data obtained in recent years, and expected in the near future, allows us to test and possibly rule out any alternative model that does not simply reduce to Λ CDM in some limit.

An idea which is particularly fascinating from the theoretical point of view is to try to explain the acceleration of the Universe by modifying General Relativity (GR) at cosmological scales, without introducing a cosmological constant. This idea started to become popular several years ago with the DGP model [2], and also underlies the intense activity of the last few years on the dRGT theory of massive gravity [3–5] (see [6, 7] for reviews), as well as on the Hassan-Rosen bigravity theory [8] (see [9] for a recent review). These studies have revealed that, even if one succeeds in the challenging task of constructing a theory of modified gravity which is theoretically consistent (e.g. ghost-free and without causality violations), one is in general still far from obtaining an interesting cosmological model. For instance, the DGP model has a self-accelerating branch [10, 11], but it succumbed to a fatal ghost-like instability which is present in this branch [12–16]. Ghost-free massive gravity has difficulties even in recovering a viable cosmological evolution at the background level, since there are no isotropic flat FRW solutions [17]. On the other hand there are open FRW solutions, which however suffer from strong coupling and ghost-like instabilities [18]. Bigravity has cosmological FRW solutions, that have been much studied recently (see e.g. [19–30]) and fall into two branches. One is a so-called infinite branch, where at the background level there is a viable self-accelerating solution; however, at the level of cosmological perturbations it suffers from fatal ghost-like instabilities both in the scalar and in the tensor sectors [31–33]. The finite branch contains instead a scalar instability at early time [31, 34, 35], which invalidates the use of linear perturbation theory and the predictivity of the theory. Tuning the Planck mass associated to the second metric, this instability can be pushed back to unobservably early times [36]. However, in this limit the theory becomes indistinguishable from Λ CDM with a graviton mass which goes to zero. It is therefore of limited interest, for the purpose of testing GR and Λ CDM against a competitor theory. Besides these two branches, in bigravity there is also a ‘non-dynamical’ branch, which also has a number of pathologies [9].

The introduction of non-local terms opens up new possibilities for building models that modify gravity in the infrared (IR), since in the IR operators such as the inverse d’Alembertian, \square^{-1} , become relevant. Even if locality is one of the basic principles of quantum field theory, at an effective level non-local terms unavoidably arise. Indeed, non-localities emerge already at a classical level, when one integrates out some fast degree of freedom to obtain an effective theory for the slow degrees of freedom, while at the quantum level loops involving light or massless particles generate non-local terms in the quantum effective action. One can therefore begin with a purely phenomenological attitude, looking for non-local terms that, added to the standard Einstein equations without cosmological

constant, produce a viable cosmological evolution. Of course, for this program to be successful, one must eventually also be able to derive such non-localities from a fundamental local theory, displaying the mechanism that generates them. However it is quite natural, and indeed probably unavoidable, to start with a purely phenomenological attitude, since identifying first non-local terms that “work”, i.e. give a viable cosmology, might give crucial hints on the underlying mechanism that generates them.

In this spirit, Deser and Woodard [37–39] proposed a non-local modification of gravity constructed adding to the Ricci scalar R in the Einstein-Hilbert action a term of the form $Rf(\Box^{-1}R)$ (see also [40–42] for related ideas). Observe that the function f is dimensionless, so this model does not involve an explicit mass scale, contrary to the non-local models that we will discuss below. There has been much activity in the literature in identifying the form that the function $f(\Box^{-1}R)$ should have in order to obtain a viable background evolution, see e.g. [43–50]. The result, $f(X) = a_1[\tanh(a_2Y + a_3Y^2 + a_4Y^3) - 1]$, where $Y = X + a_5$, and a_1, \dots, a_5 are coefficients fitted to the observed expansion history, does not look very natural. More importantly, once the function $f(X)$ is fixed in this way, one can compute the cosmological perturbations of the model and compare with the data. The result is that this model is ruled out with great statistical significance, at 7.8σ from redshift space distortions, and at 5.9σ from weak lensing [51].

From the above discussion it appears that constructing an IR modification of GR that is cosmologically viable (and still distinguishable from Λ CDM) is highly non-trivial. In the last couple of years, however, a different class of non-local models has emerged, in which the non-local terms are associated to a mass scale m (and which therefore are not of the Deser-Woodard type), and that appear to work remarkably well at the phenomenological level. The first model of this type was proposed in [52] (elaborating on previous works related to the degravitation idea [53–55], as well as on attempts at writing massive gravity in non-local form [56, 57]), and is defined by the non-local equation of motion

$$G_{\mu\nu} - \frac{m^2}{3} (g_{\mu\nu} \Box^{-1} R)^T = 8\pi G T_{\mu\nu}. \quad (1)$$

The superscript T denotes the operation of taking the transverse part of a tensor (which is itself a non-local operation), \Box is the covariant d’Alembertian computed with the curved-space metric $g_{\mu\nu}$, and its inverse \Box^{-1} is defined using the retarded Green’s function, to ensure causality. The extraction of the transverse part ensures that energy-momentum conservation is automatically satisfied. The factor $1/3$ provides a convenient normalization for the new mass parameter m . We will refer to this model as the “RT” model, where R stands for the Ricci scalar and T for the extraction of the transverse part.

A closed form for the action corresponding to eq. (1) is not known. This model is however closely related to another non-local model, subsequently proposed in [58], and defined by the action

$$S = \frac{m_{\text{Pl}}^2}{2} \int d^4x \sqrt{-g} \left[R - \frac{1}{6} m^2 R \frac{1}{\Box^2} R \right]. \quad (2)$$

We will call this the “RR” model. Computing the equations of motion derived from the action (2) and linearizing them over flat space one finds the same equations of motion as those obtained by linearizing eq. (1) over flat space. However, at the full non-linear level, or linearizing over a background different from Minkowski, the RT and RR models are

different. Thus, the RR model is in a sense the simplest non-local model of this class, at the level of the action, while the RT model can be considered as a sort of non-linear generalization of it. Of course, one can imagine different non-linear extensions of the RR model, so we just take the RR and RT models as two representative models of this class of non-local theories. It is however interesting to observe that, at the level of terms quadratic in the curvature, eq. (2) is the most general viable non-local action of this class. Indeed, at the quadratic level in the curvature, a basis is given by R^2 , $R_{\mu\nu}R^{\mu\nu}$ and $C_{\mu\nu\rho\sigma}C^{\mu\nu\rho\sigma}$, where $C^{\mu\nu\rho\sigma}$ is the Weyl tensor. Thus, beside the term $m^2R\Box^{-2}R$, at the quadratic level the most general non-local term of this class also contains terms proportional to $m_2^2R_{\mu\nu}\Box^{-2}R^{\mu\nu}$ and $m_3^2C_{\mu\nu\rho\sigma}\Box^{-2}C^{\mu\nu\rho\sigma}$, with some new mass parameters m_2, m_3 . However, as recently shown in [59], a term of the form $m_2^2R_{\mu\nu}\Box^{-2}R^{\mu\nu}$ does not provide a stable background evolution,² while a term $m_3^2C_{\mu\nu\rho\sigma}\Box^{-2}C^{\mu\nu\rho\sigma}$ does not contribute to the background evolution in FRW, but at the level of perturbations induces instabilities in the tensor sector. Thus, at the level of terms in the action which are quadratic in the curvature, only the term $m^2R\Box^{-2}R$ is phenomenologically viable. This already gives useful hints for searching the mechanism that could generate the required non-local terms. A recent discussion on how the non-local term of the RR and RT model could be generated in a fundamental local theory is given in [61, 62], where it is observed that these non-local terms corresponds to a dynamical mass generation for the conformal mode of the metric.

The study of the RR and RT models shows that they have a very appealing phenomenology. These models have no vDVZ discontinuity, and no Vainshtein mechanism is needed to restore continuity with GR [52, 58, 63]. Therefore both models, with $m = \mathcal{O}(H_0)$ as will be required by cosmology, pass without problems all solar system tests.³ In a cosmological setting, at the background level they both dynamically generate an effective dark energy and have a realistic background FRW evolution, without the need of introducing a cosmological constant [52, 58, 65]. The cosmological perturbations of these non-local models are well-behaved. This is true both in the scalar sector, as shown in [66], and in the tensor sector, as we will show in this paper (see also [59]). The study of the effect of these cosmological perturbations shows that (contrary e.g. to the Deser-Woodard model), the predictions of both the RR and RT models are consistent with CMB, supernovae, BAO and structure formation data [64, 66, 67].⁴ Further conceptual and phenomenological aspects of these models have been discussed in [68–77].

It is also interesting to observe that, when the mass parameter m is taken to be of order H_0 , non-local terms such as that in eq. (2) modify the evolution near the present cosmological epoch, but are irrelevant at early times. These non-local models can therefore be supplemented with any desired inflationary potential in the early Universe, providing standard inflation at early time, independently of the fact that they generate dark energy at late time. In particular this means that the typical values for the amplitude and tilt

²This is analogous to the fact that, in a Deser-Woodard type model, a term of the form $R^{\mu\nu}f(\Box^{-1}R_{\mu\nu})$ shows instabilities [60].

³See also app. B below, where we will discuss a related issue on the comparison with Lunar Laser Ranging, raised in [64].

⁴It should also be appreciated that a non-local model such as the RT or the RR model only introduces one new parameter m , which replaces the cosmological constant in Λ CDM. By comparison, bigravity replaces the cosmological constant by a set of 5 parameters β_n , $n = 0, \dots, 4$ and also introduces a new Planck mass associated to the second metric, and viable solutions are searched tuning this parameter space. Similarly, in the Deser-Woodard model one tunes a whole function $f(X)$.

A_s and n_s of the primordial scalar fluctuations can be obtained, in the non-local models, exactly as in Λ CDM, i.e. supplementing the model with an inflationary model at high energies. Furthermore, as first discussed in [61] (see also [76] [77] [78]), for these non-local models there is a very natural way of extending them so to obtain Starobinsky inflation at high energies. For instance, the action

$$S = \frac{m_{\text{Pl}}^2}{2} \int d^4x \sqrt{-g} \left[R + \frac{1}{6M_S^2} R \left(1 - \frac{\Lambda_S^4}{\square^2} \right) R \right], \quad (3)$$

(where $M_S \simeq 10^{13}$ GeV is the mass scale of the Starobinski model and $\Lambda_S^4 = M_S^2 m^2$) naturally interpolates between Starobinski inflation at early times and the RR model (2) at late times. Similarly, in the RT model we can combine the non-local contribution in eq. (1) with the contribution to the equations of motion coming from the R^2 term in the Starobinski model. As discussed in [61, 77] (see also [78]), at early times the non-local term is irrelevant and we recover the evolution of the Starobinski inflationary model, while at late times the local R^2 term becomes irrelevant and we recover the evolution of the non-local models.

In this paper, improving and expanding the analysis that we have presented in ref. [79] (henceforth referred as Paper I), we implement the cosmological perturbations of the two non-local models into a Boltzmann code, and perform a detailed comparison with the data as well as Bayesian parameter estimation. We can then assess the performance of the two non-local models and see if they can ‘defy’ Λ CDM, from the point of view of fitting the data. This is a level of comparison with the data, and with Λ CDM, that none of the modified gravity models discussed above have reached.

We improve on the analysis of Paper I in several respects. We use the *Planck* 2015 data, which were not yet publicly available at the time that Paper I appeared. We also include a more extended set of BAO data, including both isotropic and anisotropic BAO (while for SNe we use, as in Paper I, the JLA dataset) and we also add a comparison with growth rate data. We then perform a Bayesian model comparison between the RR, RT and Λ CDM models, using the Savage-Dickey method, embedding them in a larger model with both the non-local term and the cosmological constant. Finally, we also show that tensor perturbations in these non-local models are well-behaved, and in fact basically indistinguishable from Λ CDM.

The paper is organized as follows. In section 2 we discuss the datasets that we use and give the results obtained from cosmological parameter extraction and goodness-of-fit for the non-local models, using CMB, BAO, type Ia supernovae and H_0 measurements in different combinations, and we compare with the results obtained for Λ CDM with the same datasets. In sect. 3 we perform a Bayesian comparison between the non-local models and Λ CDM, by computing the corresponding Bayes factors using the Savage-Dickey density ratio. In sect. 4, using the best-fit values of the cosmological parameters determined by CMB, BAO and supernovae, we compare the prediction of the non-local models with structure formation and growth rate data. Tensor perturbations are discussed in sect. 5. In app. A we provide technical details about the implementation of the nonlocal models into a Boltzmann code and exhibit explicitly the equations used, while in app. B we discuss an issue on the comparison with Lunar Laser Ranging data. Our notation and conventions are the same as in Paper I [79] and in [66]. Our modified Boltzmann code is publicly available on GitHub [80].

2 Comparison with CMB, SNe and BAO data

The equations of motion of the RR and RT non-local models, both at the background and linear scalar perturbations level over FRW, have been discussed at length in [66]. For clarity we recall them briefly in app. A, in the format used in our numerical analysis. We have implemented these equations in CLASS [81], and constrained these models with observations using the Markov Chain Monte Carlo (MCMC) code Montepython [82], originally interfaced with the CLASS Boltzmann code. First of all this allows us to apply Bayesian parameter inference on the two non-local models. We can then compare the models to each other, as well as to the concordance model Λ CDM, through Bayesian model selection, given the data.

2.1 Datasets

The main datasets considered throughout this paper are the same as the ones used in the *Planck* 2015 papers, in particular in [83] for constraining the base Λ CDM model, and in [84] for constraining various phenomenological dark energy models. We adopt this choice so as to consider the most conservative data up to date, where the systematic uncertainties are mostly under control. For completeness we review them briefly in this section, referring the reader to the original papers for more detailed explanations.

CMB. We consider the likelihoods given in the recent *Planck* 2015 [85] measurements of the angular (cross-)power spectra of the CMB. This will update the analysis presented in Paper I [79], where we used the *Planck* 2013 nominal mission temperature data [86]. In particular, we take the (full-mission) lowTEB data for low multipoles ($\ell \leq 29$) and the high- ℓ Plik TT,TE,EE (cross-half-mission) ones for the high multipoles ($\ell > 29$) of the temperature and polarization auto- and cross- power spectra [84, 87].

Furthermore, since the non-local models describe a dynamical and clustering dark energy that emerges at late times [66], we also include the temperature+polarization (T+P) lensing data (where only the conservative multipole range $\ell = 40 - 400$ is used), that provide CMB constraints on late-time cosmology, and more generically allow one to break degeneracies in the primary CMB anisotropies [88, 89].

In addition to CMB data, following the *Planck* analysis [83, 84], we also include datasets from astrophysical measurements, so as to break further CMB degeneracies and reach tighter constraints on the parameter space. In particular, we consider the following datasets.

Type Ia supernovae. We consider the data from the SDSS-II/SNLS3 Joint Light-curve Analysis (JLA) [90] for SN Ia, using the complete (non-compressed) corresponding likelihoods. Combining the latter with the ones obtained from the *Planck* data allows one to put an independent constraint on the matter density fraction Ω_m , and breaks the CMB degeneracy in the $H_0 - \Omega_m$ plane. This is particularly useful in constraining the non-local models, as already pointed out in Paper I.

Baryon Acoustic Oscillations. As discussed in detail in [52, 58, 65], the RR and RT non-local models have a phantom dark energy equation of state, $w_{\text{DE}}(z) < -1$ for $z \geq 0$. Thus, the history of the growth of structures in the late universe is modified compared to a Λ CDM scenario. To further test this feature, it is useful to include different BAO scale measurements. As datasets we consider the isotropic constraints provided by 6dFGS at $z_{\text{eff}} = 0.106$ [91], SDSS-MGS DR7 at $z_{\text{eff}} = 0.15$ [92] and BOSS LOWZ at $z_{\text{eff}} = 0.32$ [93]. These data provide measurements of the acoustic-scale distance ratio $D_V(z_{\text{eff}})/r_d$. We also include the anisotropic constraints from CMASS at $z_{\text{eff}} = 0.57$ [93]. Anisotropic constraints separate the clustering effects into their longitudinal and transverse component relative to the line-of-sight. This allows one to put separate constraints on the ratios $H(z_{\text{eff}})/r_d$ and $D_A(z_{\text{eff}})/r_d$, where D_A is the angular diameter distance, and breaks the degeneracy in the D_A - H plane which arises when considering isotropic constraints. By comparison, in Paper I we only included the isotropic CMASS measurements of ref. [93], and did not consider the point provided by [92].

H_0 prior. Finally, as in Paper I, we will perform several analysis including different combinations of these datasets. First we will constrain the models given the *Planck* 2015 likelihoods only, second we will join to them the ones of JLA and BAO, and finally we will further add two different priors on H_0 provided by different local observations. These values will be taken to be $H_0 = 70.6 \pm 3.3$ [94] and $H_0 = 73.8 \pm 2.4$ [95]. In particular, we will use the value $H_0 = 73.8 \pm 2.4$ as an example of the impact of a high value of H_0 . The most recent analysis of local measurements, which appeared after this work was finished, gives $H_0 = 73.02 \pm 1.79$ [96].

2.2 Parameter space and MCMC

The datasets discussed above will be used for constraining the three statistical models that for definiteness we denote by \mathcal{M}_Λ , \mathcal{M}_{RT} and \mathcal{M}_{RR} and are associated to their respective cosmological models. For the consistency of our analysis, the initial conditions (inflationary scenario), ionization history and matter content of the universe will be chosen following the *Planck* baseline [83, 97]. In each of the non-local models we have a parameter, m^2 , that replaces the cosmological constant in Λ CDM, so the non-local models have the same number of parameters as Λ CDM. Furthermore, in the spatially flat case that we are considering, in Λ CDM the dark energy density fraction Ω_Λ can be taken as a derived parameter, fixed in terms of the other parameters by the flatness condition. Similarly, in the non-local models m^2 can be taken as a derived parameter, fixed again by the flatness condition. Thus, not only the non-local models have the same number of parameters as Λ CDM, but in fact these can be chosen so that the independent parameters are exactly the same in the non-local models and in Λ CDM, which facilitates the comparison.⁵

⁵As discussed, e.g. in [52, 58, 65], these non-local models can be formally put in a local form introducing auxiliary fields such as $U = -\square^{-1}R$ and $S = -\square^{-1}U$ (for the RR model, and similarly for the RT model, which rather involves an auxiliary four-vector field). This in principle raises the question of whether the initial conditions for these auxiliary fields are further free parameters of the model. At the conceptual level, it is important to realize that the initial conditions of these fields are actually fixed within a given model, and in principle should follow from the fundamental theory that generates these effective non-localities. Therefore, they do not correspond to extra dynamical degrees of freedom (see in particular the discussion below eq. 3.6 in [65] and refs. therein, as well as ref. [69]). From the practical point of view, this still leaves

For the neutrino masses will use the same values used in the *Planck* 2015 baseline analysis [83], i.e. two massless and a massive neutrinos, with $\sum_\nu m_\nu = 0.06$ eV, and we are fixing the effective number of neutrino species to $N_{\text{eff}} = 3.046$. In the Conclusions, following [98], we will shortly comment on the effect of the neutrino masses.

As independent cosmological parameters we take the Hubble parameter today $H_0 = 100h \text{ km s}^{-1} \text{ Mpc}^{-1}$, the physical baryon and cold dark matter density fractions today $\omega_b = \Omega_b h^2$ and $\omega_c = \Omega_c h^2$, respectively, the amplitude A_s of primordial scalar perturbations, the spectral tilt n_s and the reionization optical depth τ_{re} , so we have a 6-dimensional parameter space. We choose unbounded flat priors on all these parameters, except for τ_{re} which is taken to be bounded from below at the value of 0.01. The corresponding vector in parameter space is given by,

$$\theta = (H_0, \omega_b, \omega_c, A_s, n_s, \tau_{\text{re}}). \quad (4)$$

This parameter space will be explored via a Metropolis-Hasting algorithm. Our construction of the MCMC proceeds in two steps. The first consists in constructing a chain using wide Gaussian proposal distributions, which for all three models is taken to be centered initially on the *Planck* best fit values obtained for ΛCDM . The posterior distribution obtained then provides a new proposal distribution (covariance matrix) for a second run, where we restart from the best-fit points of the latter posterior such that the burn-in is avoided. In order to get reliable results and well shaped final distributions, we typically constructed two or three chains of 10^5 trials each and adjusted the jumping factor so as to get an acceptance rate $0.2 < r < 0.4$. The convergence of the set of chains is checked through a Gelman-Rubin convergence diagnostic. The best-fit points for each distribution are then refined by starting from the best-fit values of the (final) global run, and constructing a “cold” chain, i.e. a chain controlled by a constant temperature parameter T introduced into the MCMC acceptance probability as,

$$\alpha(\theta_n, \theta_{n+1}) = \min \left\{ 1, \left(\frac{P(\theta_{n+1})}{P(\theta_n)} \right)^{1/T} \right\}, \quad (5)$$

where θ_n is the n -th point sampled in parameter space and $P(\theta_n)$ its likelihood. Therefore the lower T is, the stronger the chain will converge towards the nearest maximum of the likelihood distribution.⁶ In our analysis we choose $T = 50$.

2.3 Results

In Table 1 we show the mean values and the corresponding χ^2 for ΛCDM , the RT model and the RR model, for different combinations of datasets. In the upper left table we only

us with the issue of making an appropriate choice of initial conditions. For the RR model we will choose the initial conditions $U(t_*) = \dot{U}(t_*) = 0$ and $S(t_*) = \dot{S}(t_*) = 0$ at some initial time t_* deep in RD (and similarly for the auxiliary fields of the RT model). As shown in sect. 4.1 of [65], the corresponding solution is an attractor in the space of solutions with different initial conditions so, within a large attraction basin, our results are actually independent of this choice. In other words, the homogeneous solutions associated to equation such as $\square U = -R$ are all decaying modes. In fact, it is just this property that selected the RR and RT model with respect to other models, such as the one featuring a term $(\square^{-1} G_{\mu\nu})^T$, originally considered in [57], for which instead the homogeneous equations have growing modes.

⁶Observe that nothing guarantees that the corresponding maximum reached is *the* global one, but starting the evaluation from the best-fit point obtained from the global run provides already representative best-fit values, sufficient for the purpose of our study.

	Planck			BAO+Planck+JLA		
Param	Λ CDM	RT	RR	Λ CDM	RT	RR
$100 \omega_b$	$2.225^{+0.016}_{-0.016}$	$2.224^{+0.016}_{-0.016}$	$2.227^{+0.016}_{-0.016}$	$2.228^{+0.014}_{-0.015}$	$2.223^{+0.014}_{-0.014}$	$2.213^{+0.014}_{-0.014}$
ω_c	$0.1194^{+0.0014}_{-0.0015}$	$0.1195^{+0.0014}_{-0.0015}$	$0.1191^{+0.0014}_{-0.0015}$	$0.119^{+0.0011}_{-0.0011}$	$0.1197^{+0.0011}_{-0.00096}$	$0.121^{+0.001}_{-0.001}$
H_0	$67.5^{+0.65}_{-0.66}$	$68.86^{+0.69}_{-0.7}$	$71.51^{+0.81}_{-0.84}$	$67.67^{+0.47}_{-0.5}$	$68.76^{+0.46}_{-0.51}$	$70.44^{+0.56}_{-0.56}$
$\ln(10^{10} A_s)$	$3.064^{+0.025}_{-0.025}$	$3.057^{+0.026}_{-0.026}$	$3.047^{+0.026}_{-0.025}$	$3.066^{+0.019}_{-0.026}$	$3.056^{+0.021}_{-0.023}$	$3.027^{+0.027}_{-0.023}$
n_s	$0.9647^{+0.0048}_{-0.0049}$	$0.9643^{+0.0049}_{-0.005}$	$0.9649^{+0.0049}_{-0.0049}$	$0.9656^{+0.0041}_{-0.0043}$	$0.9637^{+0.0039}_{-0.0041}$	$0.9601^{+0.004}_{-0.0039}$
τ_{re}	$0.0653^{+0.014}_{-0.014}$	$0.06221^{+0.014}_{-0.014}$	$0.05733^{+0.014}_{-0.014}$	$0.06678^{+0.011}_{-0.013}$	$0.0611^{+0.011}_{-0.013}$	$0.04516^{+0.014}_{-0.012}$
z_{re}	$8.752^{+1.4}_{-1.2}$	$8.442^{+1.5}_{-1.2}$	$7.932^{+1.5}_{-1.2}$	$8.893^{+1.1}_{-1.2}$	$8.359^{+1.2}_{-1.2}$	$6.707^{+1.7}_{-1.2}$
σ_8	$0.8171^{+0.0089}_{-0.0089}$	$0.8283^{+0.0092}_{-0.0096}$	$0.8487^{+0.0097}_{-0.0096}$	$0.817^{+0.0076}_{-0.0095}$	$0.8283^{+0.0085}_{-0.0093}$	$0.8443^{+0.01}_{-0.0099}$
χ^2_{min}	12943.3	12943.1	12941.7	13631.0	13631.6	13637.0
$\Delta\chi^2_{\text{min}}$	1.6	1.4	0	0	0.6	6.0

	BAO+Planck+JLA+($H_0 = 70.6$)			BAO+Planck+JLA+($H_0 = 73.8$)		
Param	Λ CDM	RT	RR	Λ CDM	RT	RR
$100 \omega_b$	$2.229^{+0.014}_{-0.015}$	$2.223^{+0.014}_{-0.014}$	$2.215^{+0.014}_{-0.014}$	$2.233^{+0.014}_{-0.014}$	$2.226^{+0.014}_{-0.014}$	$2.217^{+0.014}_{-0.014}$
ω_c	$0.1188^{+0.001}_{-0.0011}$	$0.1197^{+0.001}_{-0.0011}$	$0.1208^{+0.00099}_{-0.001}$	$0.1185^{+0.00097}_{-0.0011}$	$0.1194^{+0.001}_{-0.001}$	$0.1207^{+0.00096}_{-0.00097}$
H_0	$67.75^{+0.48}_{-0.47}$	$68.75^{+0.49}_{-0.48}$	$70.57^{+0.54}_{-0.56}$	$67.93^{+0.48}_{-0.43}$	$68.91^{+0.49}_{-0.5}$	$70.65^{+0.52}_{-0.54}$
$\log(10^{10} A_s)$	$3.069^{+0.024}_{-0.024}$	$3.056^{+0.026}_{-0.022}$	$3.03^{+0.021}_{-0.021}$	$3.077^{+0.026}_{-0.019}$	$3.061^{+0.026}_{-0.022}$	$3.031^{+0.018}_{-0.022}$
n_s	$0.9662^{+0.0042}_{-0.0042}$	$0.9637^{+0.0041}_{-0.0042}$	$0.9607^{+0.0039}_{-0.0041}$	$0.9671^{+0.0041}_{-0.0041}$	$0.9645^{+0.004}_{-0.0041}$	$0.9611^{+0.0038}_{-0.004}$
τ_{re}	$0.06883^{+0.012}_{-0.013}$	$0.06099^{+0.014}_{-0.011}$	$0.04701^{+0.011}_{-0.011}$	$0.07275^{+0.014}_{-0.01}$	$0.0641^{+0.013}_{-0.012}$	$0.04791^{+0.01}_{-0.011}$
z_{re}	$9.081^{+1.2}_{-1.1}$	$8.341^{+1.4}_{-1}$	$6.922^{+1.3}_{-1.1}$	$9.435^{+1.3}_{-0.85}$	$8.636^{+1.3}_{-1.1}$	$7.02^{+1.1}_{-1.2}$
σ_8	$0.8179^{+0.0089}_{-0.0089}$	$0.8283^{+0.0095}_{-0.0089}$	$0.8452^{+0.0085}_{-0.0086}$	$0.8197^{+0.0096}_{-0.0075}$	$0.8298^{+0.0095}_{-0.0086}$	$0.8456^{+0.0081}_{-0.0088}$
χ^2_{min}	13631.9	13631.9	13637.0	13637.5	13636.1	13638.9
$\Delta\chi^2_{\text{min}}$	0	0	5.1	1.4	0	2.8

Table 1: Parameter tables for Λ CDM and the non-local models.

use the *Planck* CMB data presented in the previous subsection, while in the upper right table we combine them with BAO and JLA supernovae. In the two lower tables we further add a prior on H_0 , namely $H_0 = 70.6 \pm 3.3$ for the lower left table and a high value, chosen for definiteness $H_0 = 73.8 \pm 2.4$, for the lower right table. Beside giving the mean values of our six fundamental parameters, we also give the corresponding derived values of σ_8 and of the effective redshift to reionization z_{re} .

In our case the models to compare among each other have the same number of parameters. Let us recall that, in this case, the standard Akaike or Bayesian information criteria [99] amount to comparing directly their respective goodness-of-fit, computing $|\Delta\chi^2_{ij}| = |\chi^2_{\text{min},i} - \chi^2_{\text{min},j}|$ between models \mathcal{M}_i and \mathcal{M}_j . The larger the value of $|\Delta\chi^2_{ij}|$, the higher the evidence for the model with smaller goodness-of-fit. A difference $|\Delta\chi^2_{ij}| \leq 2$ implies statistically equivalence between the two models compared, while $|\Delta\chi^2_{ij}| \gtrsim 2$ suggests “weak evidence”, and $|\Delta\chi^2_{ij}| \gtrsim 6$ indicates “strong evidence”. We will come back to the model comparison in a fully Bayesian manner in Section 3, and for the moment we stick to the comparison of the χ^2 .

Comparing with Tables 1-3 of Paper I, we see that the present results confirm and

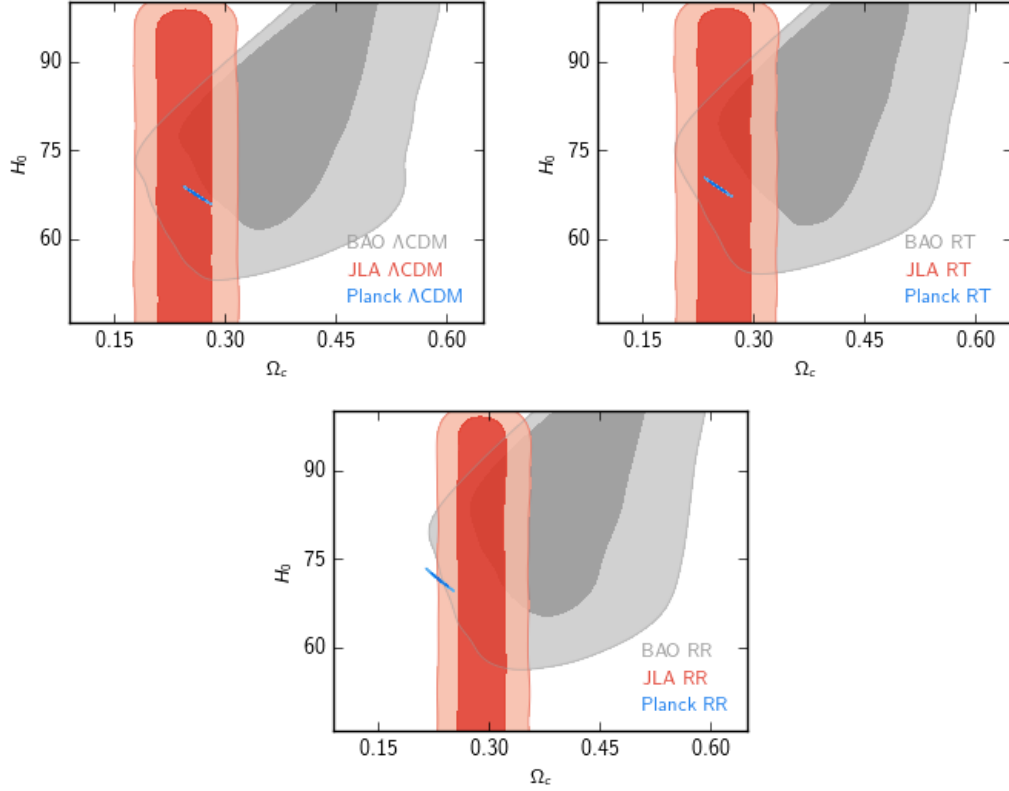


Figure 1: The separate 1σ and 2σ contours for CMB, BAO and SNe in the plane (H_0, Ω_c) for the models Λ CDM, RT and RR.

reinforce the trends already observed in Paper I. In particular, with CMB only, we see from the values of the χ^2 that the three models perform in a very comparable manner, with the RR model being slightly better, but at a level which is not statistically significant. However, once we add BAO and SNe, Λ CDM and the RT model remain statistically indistinguishable, while the RR model is significantly disfavored. This trend is strongly reinforced in the present analysis compared to Paper I, since now $\chi^2_{\min, \text{RR}} - \chi^2_{\min, \Lambda\text{CDM}} \simeq 6.0$, while in Paper I we found $\chi^2_{\min, \text{RR}} - \chi^2_{\min, \Lambda\text{CDM}} \simeq 3.2$. Just as in Paper I, the origin of this result can be understood looking at the two-dimensional marginalized likelihoods in the $H_0 - \Omega_c$ plane, shown in Fig. 1 (to be compared with Fig. 1 of Paper I). We see that for the RT model and for Λ CDM the separate contours obtained from CMB, BAO and JLA supernovae agree very well, while for the RR model there is a slight tension between CMB and JLA. We emphasize, though, that even this ‘tension’ is just at the 2σ level. In other words even the RR model, by itself, fits the data at a fully acceptable level. However, in the comparison with the RT model and with Λ CDM, it clearly performs less well, when no prior on H_0 is imposed (and assuming the Planck baseline value for the neutrino masses, see the Conclusions). Observe that the shift in the BAO contour, with respect to Paper I, is mostly due to the inclusion of the SDSS-MGS DR7 point at $z_{\text{eff}} = 0.15$.

Concerning the mean values, we see from the table that the most significant difference is in H_0 , with the non-local model (and particularly the RR model) favoring slightly higher

Data	RT	RR
Planck	$5.17(4) \times 10^{-2}$	$9.35(7) \times 10^{-3}$
BAO+Planck+JLA	$5.15(4) \times 10^{-2}$	$9.21(7) \times 10^{-3}$
BAO+Planck+JLA+($H_0 = 70.6$)	$5.15(4) \times 10^{-2}$	$9.22(7) \times 10^{-3}$
BAO+Planck+JLA+($H_0 = 73.8$)	$5.17(4) \times 10^{-2}$	$9.24(7) \times 10^{-3}$

Table 2: Mean values for $\gamma = m^2/(9H_0^2)$.

values of H_0 , compared to Λ CDM. This confirms again a result obtained in Paper I. So, not surprisingly, adding as a prior the value of H_0 from local measurements goes in the direction of favoring the non-local models. Actually, with the value $H_0 = 70.6 \pm 3.3$ given in [94] the situation changes little, with Λ CDM and the RT model still statistically equivalent, and the RR model disfavored, while a higher value, such as $H_0 = 73.8 \pm 2.4$, would bring back the RR model to values of $\Delta\chi^2$ only weakly disfavored. In all cases, Λ CDM and the RT model are statistically indistinguishable, with differences in favor of one or the other which depend on the datasets and priors used.

From the above results we can also obtain the mean value for the derived parameter m^2 of the non-local models. Actually, it is convenient to use the dimensionless quantity

$$\gamma = \frac{m^2}{9H_0^2}, \quad (6)$$

that enters in the study of the cosmological evolution [52, 58]. Based on the values of the parameters given in Table 1 we find, for γ , the values given in Table 2.

In Fig. 2 we show the triangle plot for the case Planck+BAO+JLA plus a prior $H_0 = 70.6$, while in Fig. 3 we plot the 2-dimensional likelihood in the plane $\sigma_8 - \Omega_m$. Finally, having also determined the best fit values of the cosmological parameters, it is interesting to display explicitly how the CMB data are fitted by the three models. We use for definiteness the best fit determined from BAO+JLA+Planck. In Figs. 4-7 we show, for the three models, the fit to the temperature power spectrum, the EE spectrum, the lensed BB spectrum and the lensing potential, respectively.

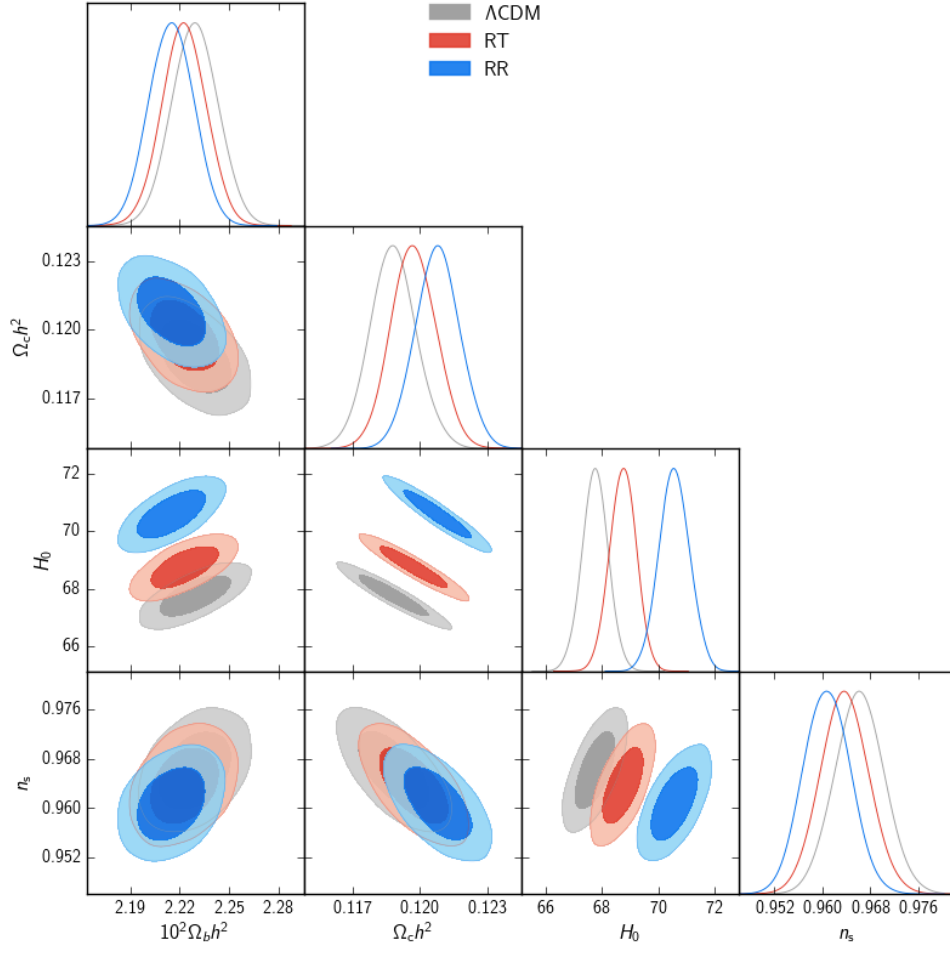


Figure 2: Triangle plot for Planck+BAO+JLA+($H_0 = 70.6$).

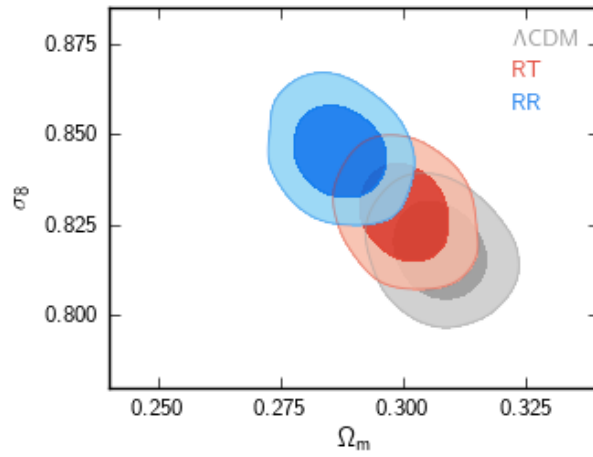


Figure 3: $\sigma_8 - \Omega_m$ contour plot for Planck+BAO+JLA+($H_0 = 70.6$).

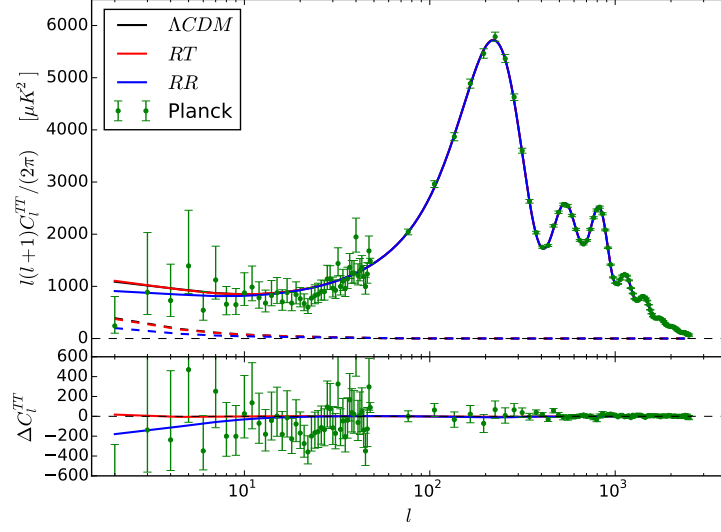


Figure 4: Upper plot: temperature power spectrum (thick), and the separate contribution from the late ISW contributions (dashed), for Λ CDM (black), RT (red) and RR (blue), using the best fit values of the parameters determined from BAO+JLA+Planck. The black and red lines are indistinguishable on this scale. The lower plot shows the residuals for Λ CDM and difference of RT (red) and RR (blue) with respect to Λ CDM. Data points are from Planck 2015 [83] (green bars). Error bars correspond to $\pm 1\sigma$ uncertainty.

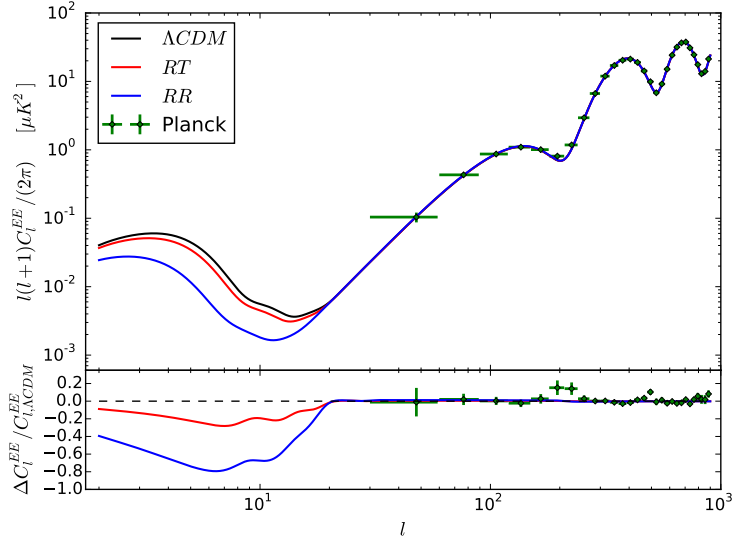


Figure 5: Similar, for the EE spectra except that the separate late ISW contributions are not shown. The lower plot shows the *relative* difference of RT (red) and RR (blue) with respect to Λ CDM.

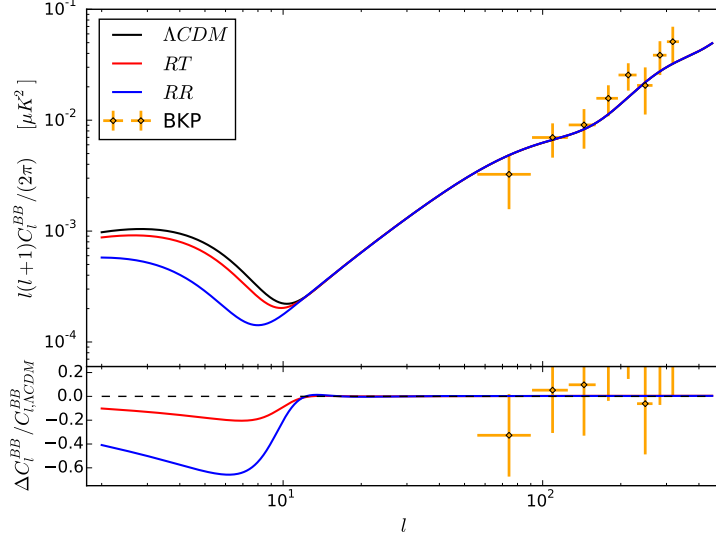


Figure 6: As in Fig. 5 for the lensed BB spectra the data points are from the joint BICEP2+Keck+Planck [100] (orange crosses).

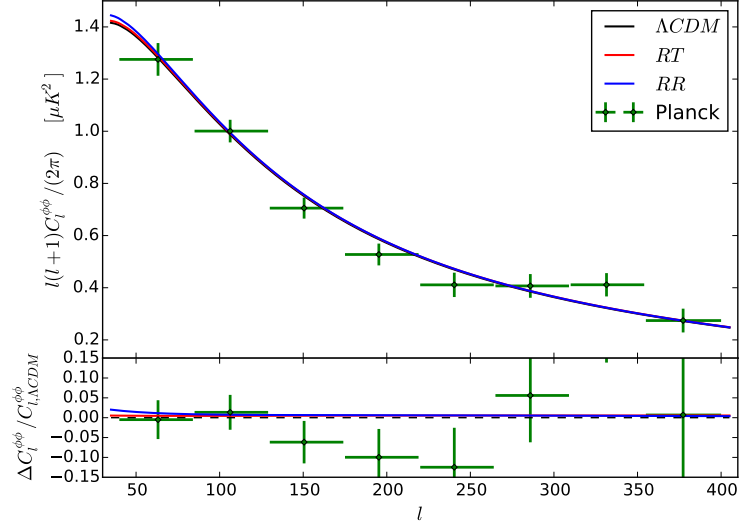


Figure 7: The lensed $\phi\phi$ spectra (lensing potential).

3 Bayesian model comparison

3.1 The Bayes factor

In the previous section, in order to compare the performance of the three models, we have used the differences $\Delta\chi^2_{ij}$. This method is not genuinely Bayesian, and only leads to sensible results if a number of assumptions are satisfied by the posterior distributions under interest (e.g. large number of data points, weak skewness, etc), so it is possible that the results drawn from it are biased. A more accurate method for comparing models, which is fully Bayesian, is based on Bayes factors [101]. In this section we use Bayes factors to compare the three models of interest among each other, given the *Planck*+BAO+JLA data. Starting from Bayes theorem,

$$P(\theta|d, \mathcal{M}) = \frac{P(d|\theta, \mathcal{M})P(\theta|\mathcal{M})}{P(d|\mathcal{M})}, \quad (7)$$

which says that the posterior distribution, i.e. the probability of the parameters θ given the data d and the model \mathcal{M} , $P(\theta|d, \mathcal{M})$, equals the product of the likelihood function $P(d|\theta, \mathcal{M})$ by the prior $P(\theta|\mathcal{M})$, divided by the evidence $P(d|\mathcal{M})$ (marginal likelihood). General Bayesian model comparison is based on the model probability $P(\mathcal{M}|d)$. This number is however difficult to interpret in absolute terms, and instead one considers relative model probabilities. Using again Bayes theorem we can express the relative model probability for two models \mathcal{M}_i and \mathcal{M}_j as

$$\frac{P(\mathcal{M}_i|d)}{P(\mathcal{M}_j|d)} = \frac{P(d|\mathcal{M}_i)}{P(d|\mathcal{M}_j)} \frac{P(\mathcal{M}_i)}{P(\mathcal{M}_j)} \equiv B_{ij} \frac{P(\mathcal{M}_i)}{P(\mathcal{M}_j)}, \quad (8)$$

i.e. up to the prior model probabilities $P(\mathcal{M})$ the relative model probabilities are just given by the Bayes factors B_{ij} which are defined as the ratio of the evidences computed within the two models \mathcal{M}_i and \mathcal{M}_j . The Bayes factors correspond to betting odds, their numerical value is conventionally translated into statements about the evidence of model i with respect to model j using the Jeffreys' scale. Model i is favored with respect to model j if $B_{ij} > 1$. For $1 < B_{ij} < 3$ the evidence is deemed 'weak', for $3 < B_{ij} < 20$ is 'definite', for $20 < B_{ij} < 150$ is 'strong', and for $B_{ij} > 150$ is 'very strong'.⁷

For multi-parameter models the computation of the respective evidences is in general numerically expensive. However when the models to compare with each other can be nested together, i.e. can be embedded in a larger model that reduces to one model in a limit, and to the other model in another limit, this task is made easier by the Savage-Dickey density ratio (SDDR) method. In our context this nesting is possible and the SDDR can be applied.

3.2 Model nesting and Savage-Dickey density ratio

Our task is to compute the Bayes factors $B_{\Lambda i}$ constructed out of the evidence computed assuming the Λ CDM model, our null hypothesis, and one of the two non-local models of interest ($i = \text{RT, RR}$), the alternative hypothesis, for a given set of data. Computing the

⁷Some subtleties may appear in the use of the Jeffreys' scale when comparing models with a different number of parameters [102], which however is not our case.

evidence for a model is generally quite hard as it requires a high-dimensional integration. Marginalising Eq. (7) over θ we find immediately that

$$P(d|\mathcal{M}) = \int d\theta P(d|\theta, \mathcal{M})P(\theta|\mathcal{M}). \quad (9)$$

If however two models are nested, it becomes possible to use the SDDR instead to find directly the Bayes factor between the models. In order to be able to exploit the SDDR we consider the extended model constructed from each of the non-local models by adding a cosmological constant. For the model RT the equation of motion (1) is then modified into

$$G_{\mu\nu} - \frac{m^2}{3} (g_{\mu\nu} \square^{-1} R)^T + g_{\mu\nu} \Lambda = 8\pi G T_{\mu\nu}, \quad (10)$$

while for the model RR the action (2) becomes

$$S = \frac{m_{\text{Pl}}^2}{2} \int d^4x \sqrt{-g} \left[R - 2\Lambda - \frac{1}{6} m^2 R \frac{1}{\square^2} R \right]. \quad (11)$$

The model (10) reduces to Λ CDM in the limit $m^2 \rightarrow 0$, and to the RT model in the limit $\Lambda \rightarrow 0$, and similarly the model (11) reduces in these limits to Λ CDM and to the RR model, respectively. Of course, the models (10) and (11) could be interesting in their own right. However, we will only use them as a tool for comparing the RR and RT models to Λ CDM.

These two models for gravity allow us to construct two associated statistical models, denoted $\mathcal{M}_{\Lambda+\text{RT}}$ and $\mathcal{M}_{\Lambda+\text{RR}}$ respectively, which can be constrained by the data with the same method described in the previous section. The only difference with respect to the analysis performed previously is that now the total dark energy component is a mixture of the one induced by the cosmological constant, Ω_Λ , and the one induced by the non-local modification of gravity Ω_{X_i} , which depends on the mass parameter m . These models therefore have an extended parameter space that we take to be spanned by

$$\tilde{\theta} = (H_0, \omega_b, A_s, n_s, \tau_{\text{re}}, \Omega_\Lambda, \Omega_{X_i}). \quad (12)$$

Observe that now the physical dark matter density fraction ω_c is taken as a derived parameter, and instead we vary the density fractions of the two types of dark energy. We choose to proceed in this way in order to keep maximal control on the choice of prior for the latter, since this condition can be very important when using the SDDR method [103] that we discuss now.

Consider the nested model $\mathcal{M}_{\Lambda+i}$ (where $i = \text{RT}$ or RR). Since this model is a nesting of \mathcal{M}_Λ and \mathcal{M}_i , we assume (by continuity) that its likelihood function taken at $\Omega_\Lambda = 0$ reproduces the one of \mathcal{M}_i , and equivalently putting $\Omega_{X_i} = 0$ reproduces the one of \mathcal{M}_Λ ,

$$P(d|\bar{\theta}_\Lambda, \Omega_\Lambda = 0, \mathcal{M}_{\Lambda+i}) = P(d|\bar{\theta}_\Lambda, \mathcal{M}_i), \quad (13)$$

$$P(d|\bar{\theta}_{X_i}, \Omega_{X_i} = 0, \mathcal{M}_{\Lambda+i}) = P(d|\bar{\theta}_{X_i}, \mathcal{M}_\Lambda), \quad (14)$$

where $\bar{\theta}_j \equiv \tilde{\theta} \setminus \{\Omega_j\}$, and that the same conditions hold on the prior. We can therefore

write

$$\begin{aligned}
P(d|\mathcal{M}_\Lambda) &= \int d\bar{\theta} P(d|\bar{\theta}, \mathcal{M}_\Lambda) P(\bar{\theta}|\mathcal{M}_\Lambda) \\
&= \int d\bar{\theta} P(d|\bar{\theta}_{X_i}, \Omega_{X_i} = 0, \mathcal{M}_{\Lambda+i}) P(\bar{\theta}_{X_i}, \Omega_{X_i} = 0|\mathcal{M}_{\Lambda+i}) \\
&= P(d|\Omega_{X_i} = 0, \mathcal{M}_{\Lambda+i}).
\end{aligned} \tag{15}$$

Applying now Bayes theorem to the third equality,

$$P(d|\Omega_{X_i} = 0, \mathcal{M}_{\Lambda+i}) = \frac{P(\Omega_{X_i} = 0|d, \mathcal{M}_{\Lambda+i}) P(d|\mathcal{M}_{\Lambda+i})}{P(\Omega_{X_i} = 0|\mathcal{M}_{\Lambda+i})}, \tag{16}$$

leads to the SDDR,

$$B_{\Lambda(\Lambda+i)} \equiv \frac{P(d|\mathcal{M}_\Lambda)}{P(d|\mathcal{M}_{\Lambda+i})} = \frac{P(\Omega_{X_i} = 0|d, \mathcal{M}_{\Lambda+i})}{P(\Omega_{X_i} = 0|\mathcal{M}_{\Lambda+i})}, \tag{17}$$

which tells us that the Bayes factor of the comparison between the nested model $\mathcal{M}_{\Lambda+i}$ and its sub-model \mathcal{M}_Λ is equal to the ratio of the marginalized one-dimensional posterior distribution and marginalized one-dimensional prior of Ω_{X_i} obtained from the extended model, evaluated at the point where the simpler model is nested inside the extended model (i.e. at $\Omega_{X_i} = 0$). Intuitively this can be understood as a measure of the amount by which the prior evolved into the posterior through knowledge update given new data. If the posterior grows at $\Omega_{X_i} = 0$ this means that this value tends to be preferred by the data, i.e. \mathcal{M}_Λ is preferred, and the Bayes factor increases. Conversely, if the posterior decreases, a non-zero value of Ω_{X_i} is preferred.

The procedure outlined above allows us to compare \mathcal{M}_Λ (or equivalently \mathcal{M}_i) with the nested model $\mathcal{M}_{\Lambda+i}$. However, we eventually want to compare directly \mathcal{M}_Λ with \mathcal{M}_i . This can be done by comparing both to the extended model,

$$\begin{aligned}
B_{\Lambda i} &\equiv \frac{P(d|\mathcal{M}_\Lambda)}{P(d|\mathcal{M}_i)} = \frac{P(d|\mathcal{M}_\Lambda)}{P(d|\mathcal{M}_{\Lambda+i})} \frac{P(d|\mathcal{M}_{\Lambda+i})}{P(d|\mathcal{M}_i)} \\
&= \frac{B_{\Lambda(\Lambda+i)}}{B_{i(\Lambda+i)}} = \frac{P(\Omega_{X_i} = 0|d, \mathcal{M}_{\Lambda+i})}{P(\Omega_\Lambda = 0|d, \mathcal{M}_{\Lambda+i})}.
\end{aligned} \tag{18}$$

The expression involving the ratio of the Bayes factors always holds, but in the last identity we have assumed the equality of the marginalized one-dimensional prior, $P(\Omega_{X_i} = 0|\mathcal{M}_{\Lambda+i}) = P(\Omega_\Lambda = 0|\mathcal{M}_{\Lambda+i})$. This shows how the SDDR method greatly simplifies the computation of the Bayes factor for the comparison of two models that can be nested together. Indeed, one only needs to know the final posterior distribution of the latter, for some set of data, which can be obtained by standard MCMC methods as the one we use (although occasionally multiple chains at different temperatures may be needed, see e.g. appendix A of [104] for an example in a different context). Then one has just to take the ratio of two numbers, the marginalized 1d posterior of Ω_Λ evaluated at $\Omega_\Lambda = 0$, and the same for Ω_{X_i} , within the extended model. If the priors do not cancel then one just computes the actual Bayes factors between the extended and the two nested models and uses their ratio.

	BAO+Planck+JLA			BAO+Planck+JLA+($H_0 = 73.8 \pm 2.4$)	
Param	RT + Λ	RR + Λ	Param	RT + Λ	RR + Λ
100 ω_b	$2.225^{+0.014}_{-0.014}$	$2.225^{+0.015}_{-0.015}$	100 ω_b	$2.23^{+0.013}_{-0.014}$	$2.227^{+0.015}_{-0.015}$
H_0	$68.18^{+0.65}_{-0.68}$	$68.28^{+0.79}_{-1}$	H_0	$68.62^{+0.64}_{-0.58}$	$69.08^{+0.92}_{-1}$
$\log(10^{10}A_s)$	$3.059^{+0.017}_{-0.02}$	$3.061^{+0.026}_{-0.025}$	$\log(10^{10}A_s)$	$3.066^{+0.017}_{-0.017}$	$3.059^{+0.028}_{-0.026}$
n_s	$0.9646^{+0.004}_{-0.0041}$	$0.9646^{+0.0044}_{-0.0044}$	n_s	$0.9656^{+0.0039}_{-0.0039}$	$0.9649^{+0.0045}_{-0.0045}$
τ_{re}	$0.06323^{+0.0096}_{-0.011}$	$0.06379^{+0.014}_{-0.013}$	τ_{re}	$0.06716^{+0.0092}_{-0.0092}$	$0.06328^{+0.015}_{-0.014}$
Ω_Λ	$0.357^{+0.35}_{-0.33}$	$0.538^{+0.24}_{-0.17}$	Ω_Λ	$0.2445^{+0.13}_{-0.44}$	$0.3993^{+0.21}_{-0.22}$
Ω_X	$0.3368^{+0.34}_{-0.35}$	$0.1566^{+0.17}_{-0.24}$	Ω_X	$0.454^{+0.42}_{-0.17}$	$0.3025^{+0.22}_{-0.21}$
Ω_{cdm}	$0.2569^{+0.0025}_{-0.023}$	$0.2561^{+0.0076}_{-0.0072}$	Ω_{cdm}	$0.2527^{+0.0035}_{-0.022}$	$0.2501^{+0.0052}_{-0.03}$
z_{re}	8.568^{+1}_{-1}	$8.604^{+1.4}_{-1.2}$	z_{re}	$8.935^{+0.93}_{-0.86}$	$8.537^{+1.5}_{-1.3}$
σ_8	$0.8217^{+0.033}_{-0.0012}$	$0.8238^{+0.011}_{-0.012}$	σ_8	$0.8257^{+0.036}_{-0.0017}$	$0.8306^{+0.018}_{-0.035}$
χ^2_{min}	13631.0	13630.8	χ^2_{min}	13636.1	13635.4
$B_{\Lambda i}$	1.02	22.67	$B_{\Lambda i}$	0.39	2.38

Table 3: Best fit values and Bayes factors for the nested models. Left: using BAO+Planck+JLA data. Right: adding also $H_0 = 73.8 \pm 2.4$.

The last point that we need to discuss is the choice that we made for the parameters to vary for constructing the posterior distribution of the nested model, and their respective prior distribution. In Bayesian inference the choice of a (subjective) prior is part of the approach and should be chosen with care. In order to have sensible results, one needs to provide the least informative prior given by the current knowledge, before seeing the data. Since in the nested model we have two types of dark energies, Ω_{X_i} and Ω_Λ , the flatness condition reads $\Omega_{X_i} + \Omega_\Lambda + \Omega_c = 1 - \Omega_R$. In order to treat both type of dark energies on the same footing, we choose to vary both Ω_Λ and Ω_{X_i} and to take ω_c as derived. For the same reason, we prefer to impose a flat prior on Ω_Λ and Ω_{X_i} rather than on, say, Λ and m^2 . It is natural to assume that the density fractions vary in the interval $[0, 1]$. To avoid boundary effects that could affect the determination of the SDDR, we actually extend the allowed range of values to

$$\Omega_\Lambda \in [-0.2, 1.2], \quad \Omega_{X_i} \in [-0.2, 1.2], \quad (19)$$

with a uniform prior and we then remove values that lie outside $[0, 1]$ from the chain. The prior for both parameter being the same, the last equality in (18) therefore make sense, and this formula can be applied in our case.

We can now compute the likelihoods of the nested model, perform the corresponding parameter estimation, and compute the Bayes factors. The results are shown in Table 3.

Not surprisingly, the mean values for the nested model $\mathcal{M}_{\Lambda+\text{RT}}$ are always intermediate between the values obtained for \mathcal{M}_Λ and those for \mathcal{M}_{RT} , compare with Table 1, and the same for the nested model $\mathcal{M}_{\Lambda+\text{RR}}$. The conclusions drawn from the values of the Bayes factors are fully consistent with that obtained in Section 2.3 simply using $\Delta\chi^2_{ij}$. Namely, without a prior on H_0 , the RT model and ΛCDM are statistically indistinguishable while, on the Jeffreys' scale, the evidence of ΛCDM against the RR model is on the border between 'definite' and 'strong'. Adding a high prior on H_0 favors the non-local models, so that the RT models become slightly favored with respect to ΛCDM , with

$B_{\text{RT},\Lambda} \simeq 1/0.39 \simeq 2.6$, and the RR model is now only slightly disfavored with respect to ΛCDM , by a similar amount, $B_{\Lambda,\text{RR}} \simeq 2.4$, in both cases not very significant.

Also the evidence of the RT model with respect to the RR model can be computed easily, with $B_{\text{RT,RR}} = B_{\text{RT},\Lambda}/B_{\text{RR},\Lambda} = B_{\Lambda,\text{RR}}/B_{\Lambda,\text{RT}}$ ranging from a value $22.67/1.02 \simeq 22.2$ using BAO+Planck+JLA data, to a value $2.38/0.39 \simeq 6.1$ when adding the high prior on the Hubble constant, $H_0 = 73.8 \pm 2.4$.

4 Growth rate data and modified structure formation

Redshift-space distortions (RSD) probe the velocity field, and are usually expressed as a constraint on the combination $f\sigma_8$, where $f \equiv d \ln \sigma_8 / d \ln a$ is the growth function [84], and σ_8 is the variance of the linear matter power spectrum in a radius of 8 Mpc today. The translation to $f\sigma_8$ requires a fiducial model, which is taken to be ΛCDM . This could lead to problems when using the data for constraining modified gravity models, especially if they exhibit a significant scale-dependence. However, the linear matter power spectrum of the RT and RR models look very much like that of ΛCDM on the relevant scales, as can be seen from Figs. 8 and 9, which show the linear matter power spectra at redshifts $z = 0$ and $z = 1$, respectively. For this reason we expect that we can use the RSD data to provide at least a rough test of our non-local models. Furthermore, since RSD are mostly degenerated with the Alcock-Paczynski (AP) effect, measurements of $f\sigma_8$ are often combined with measurements of the angular diameter distance D_A and the Hubble parameter H at the corresponding effective redshift, or some combination of the latter. Whenever this is the case, we marginalize over these measurements and consider only the ones associated to $f\sigma_8$. From this point of view, the treatment which follows is only illustrative of the constraints partially imposed by RSD on the models under consideration. Of course, a more thorough study would consist in including the RSD datasets into a fit performed with MCMC techniques, similarly to what we performed in the previous section for CMB, SNe and BAO data. For the purpose of our study, we fix the best-fit values of the models from the *Planck*+BAO+JLA data and check if the predicted growth of perturbations agrees with the marginalized $f\sigma_8$ data. The result and the corresponding data used are shown in Fig. 10.⁸ We find that the corresponding χ^2 obtained from the data in Fig. 10 for ΛCDM , the RT and RR models are

$$\chi_{\text{min},\Lambda\text{CDM}}^2 = 3.9, \quad \chi_{\text{min,RT}}^2 = 4.7, \quad \chi_{\text{min,RR}}^2 = 6.5. \quad (20)$$

We see that the non-local models generically predict a larger growth rate, although for the RT model the difference with ΛCDM is again not statistically significant. Once again, it should be stressed that this result is specific to our choice of neutrino masses. It is known (see e.g. [83]) that increasing the value of the neutrino masses has the effect of decreasing both σ_8 and H_0 . In ΛCDM a decrease in H_0 would increase the tension with local measurements of H_0 . However, we have seen that the RR model predicts a higher value of H_0 , so it can in principle be consistent with local measurement of H_0 even for

⁸Notice that the data point of SDSS LRG analyzed by [105] is taken to be the resulting data from their “full fit” since other fits are explicitly based on ΛCDM -dependent priors such as the ones imposed on $\sigma_{8,\text{nl}}$ (which is not applicable in our case, see $\Omega_m - \sigma_8$ contours in fig. 3) or neglecting the AP effect which is motivated by the fit using the fiducial cosmology.

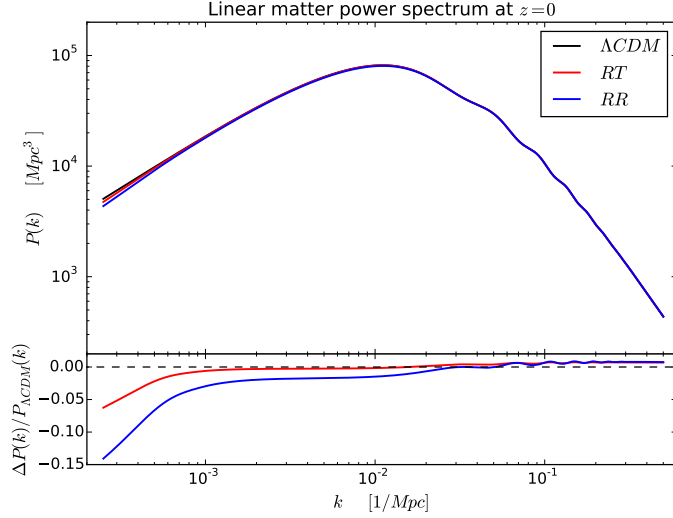


Figure 8: Upper: linear matter power spectrum for Λ CDM (black), RT (red) and RR (blue). Lower: relative difference of RT (red) and RR (blue), with respect to Λ CDM. These plots give the linear power spectrum evaluated at redshift $z = 0$.

larger neutrino masses. In turn such larger neutrino masses would have the effect of decreasing σ_8 , which could render again the RR model fully competitive with Λ CDM, also from the point of view of fitting structure formation data. The corresponding analysis will be presented in [98].

We have also computed several other quantities which are relevant for the comparison with structure formation (in particular for assessing the possibility of discriminating the three models with future observations), using for each model the respective mean values of the parameters obtained from *Planck*+BAO+JLA data. When studying structure formation and lensing, several indicators are used in the literature for parametrizing deviations from Λ CDM. We define the Bardeen potentials Ψ and Φ according to the sign convention

$$ds^2 = -(1 + 2\Psi)dt^2 + a^2(t)(1 + 2\Phi)\delta_{ij}dx^i dx^j. \quad (21)$$

Then, two useful quantities for parametrizing deviations from GR are the functions $\mu(k, z)$ [112] and $\Sigma(k, z)$ [113], which are defined through

$$\Psi = \mu(k, z)\Psi_{\text{GR}}, \quad (22)$$

$$\Psi - \Phi = \Sigma(k, z)(\Psi - \Phi)_{\text{GR}}, \quad (23)$$

where the subscript denotes the quantities computed in GR, assuming a Λ CDM model. The advantage of this parametrization is that it neatly separates the modifications to the motion of non-relativistic particles, which is described by μ , from the modification to light propagation, which is encoded in Σ . The functions μ and Σ depend on the comoving momentum k and on red-shift z . We compute the time-evolution of Ψ using the mean values for the non-local models and Ψ_{GR} using the mean values for Λ CDM.⁹ Often in the

⁹Observe that in [66], where a similar analysis (based however on Λ CDM mean values for the evolution

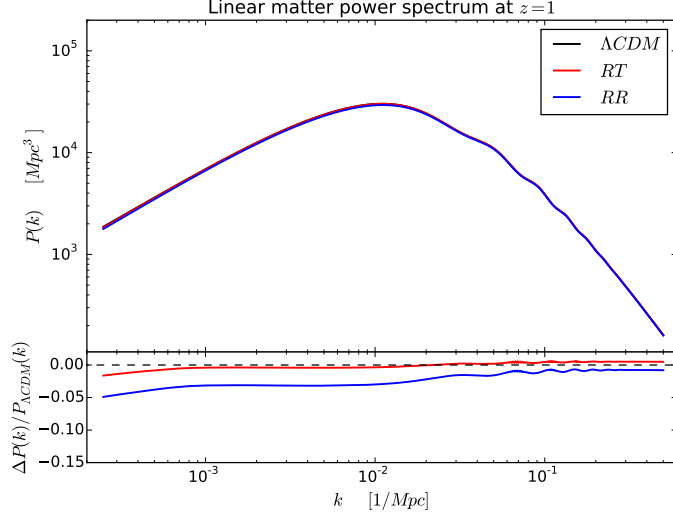


Figure 9: As in Fig. 8, for redshift $z = 1$.

literature, e.g. in the Planck DE paper [84], a slightly different convention is used:

$$-k^2\Psi = 4\pi G a^2 \mu_P(k, z) \rho \Delta, \quad (24)$$

$$-k^2(\Psi - \Phi) = 8\pi G a^2 \Sigma_P(k, z) \rho \Delta. \quad (25)$$

(The subscript P in μ_P and Σ_P indicates that these definitions agree with those of the Planck analysis [84].) In these definitions we compare the actual value of the gravitational potentials Φ and Ψ to the value expected in GR due to the matter and radiation perturbations $\rho\Delta = \rho_m\Delta_m + \rho_r\Delta_r$, where Δ is the comoving density perturbation. In Fig. 11 and 12 we plot $\mu_P(k, z)$ and $\Sigma_P(k, z)$ for two representative momenta, as a function of z , for the three models. We see that, for the RT model, the deviations from Λ CDM are tiny, below one percent, while for the RR model they reach a maximum value up to 3 – 4% at the typical redshifts, say, $z \gtrsim 0.5$ relevant for observation of structure formation.

An alternative useful pair of indicators is provided by the slip function and the effective Newton constant [49, 51, 113]. The slip function is given by¹⁰

$$\eta_P(k, z) = -\frac{\Phi}{\Psi}, \quad (26)$$

This quantity is equal to one in Λ CDM when the anisotropic stress is negligible. The effective Newton constant G_{eff} is defined as the constant that enters the Poisson equation for Φ (see sect. 4.3 and app. A of [66] for its explicit expression in the RR and RT models). Observe that $G_{\text{eff}}(k, z)/G$ is the same as the function $Q(k, z)$ of [84]. We plot these quantities in Figs. 13 and 14, while in Fig. 15 we show the combination $2[\mu_P(k, z) -$

of both Ψ and Ψ_{GR}) had been carried out, we used a different convention for the definition of μ and Σ , writing $\Psi = (1 + \mu)\Psi_{\text{GR}}$ and $(\Psi - \Phi) = (1 + \Sigma)(\Psi - \Phi)_{\text{GR}}$, respectively. With the definitions (22) and (23) the GR limit corresponds to $\mu = \Sigma = 1$ rather than $\mu = \Sigma = 0$.

¹⁰Again the subscript P means that we use the Planck conventions. Another common definition, that we used in [66], is $\eta = (\Phi + \Psi)/\Phi$.

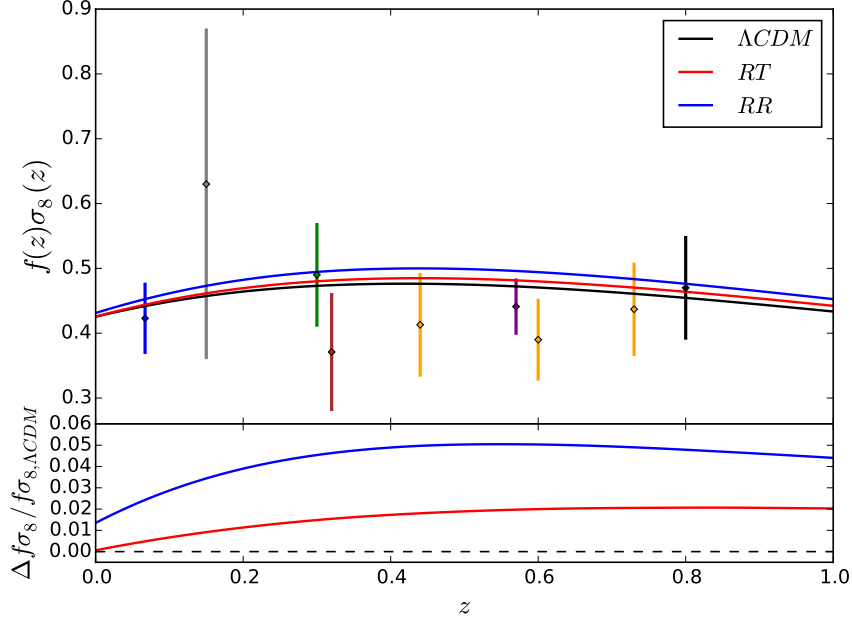


Figure 10: Upper panel: Growth rate computed in Λ CDM (black), RT (red) and RR (blue) constrained with various data points: 6dF GRS [106] (blue), SDSS LRG [105] (grey), SDSS MGS [107] (green), BOSS LOWZ [108] (brown), BOSS CMASS [109] (purple), WiggleZ [110] (orange), VIPERS [111] (black). Lower panel: Relative difference of the growth rate computed in RT (red) and in RR (blue) with respect to Λ CDM.

$1] + [\eta_P(k, z) - 1]$, which is the one used in the Planck 2015 dark energy paper [84]. Finally, we show the difference $\Psi - \Phi$ both at large and small scales in Fig. 16.

What we learn from these figures is that, at the typical redshifts of interest for the comparison with observations of structure formation, say $z \gtrsim 0.5$, as far as structure formation is concerned the RT model differs from Λ CDM at a level of at most 1%, which is quite small compared to existing experimental uncertainties, but could be observable with future missions such as Euclid [114]. In contrast, the RR model shows differences which, at $z \simeq 0.5$, can be as large as 6%, as for instance for the combination shown in Fig. 15.¹¹ As we see from Fig. 14, these deviations go in the direction of producing a stronger effective Newton constant, and therefore more structures. This is the main reason why the growth displayed in Fig. 10 is stronger for the non-local models than for Λ CDM. A further reason is that, for the same values of the cosmological parameters, the expansion rate in the RT and RR model is generically lower than in Λ CDM as can be seen in Fig. 17, and this contributes (although to a smaller level) also to the enhancement of clustering.¹²

¹¹By comparison, the Deser-Woodard model has been ruled out because μ , near $z = 0.5$, differs from the GR value by about 60% and Σ by about 20%, see Fig. 1 of [51].

¹²This fact was also noticed in [64] where N-body simulations revealed deeper gravitational potential wells and massive haloes slightly more abundant and concentrated.

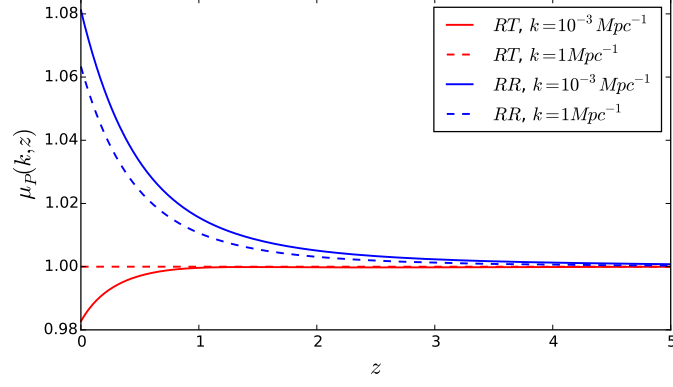


Figure 11: Deviations of the non-local models from Λ CDM for the gravitational potential parametrized by μ_P .

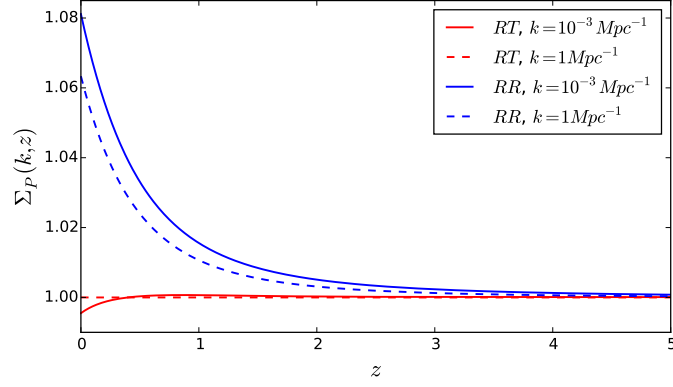


Figure 12: Deviations of the non-local models from Λ CDM for the gravitational potential parametrized by Σ_P .

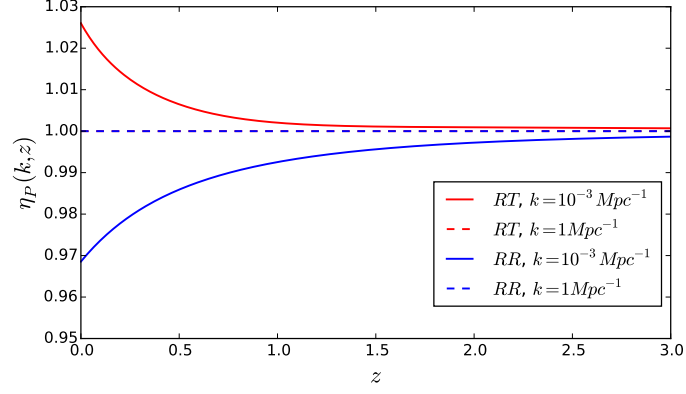


Figure 13: The gravitational slip η_P for the non-local models.

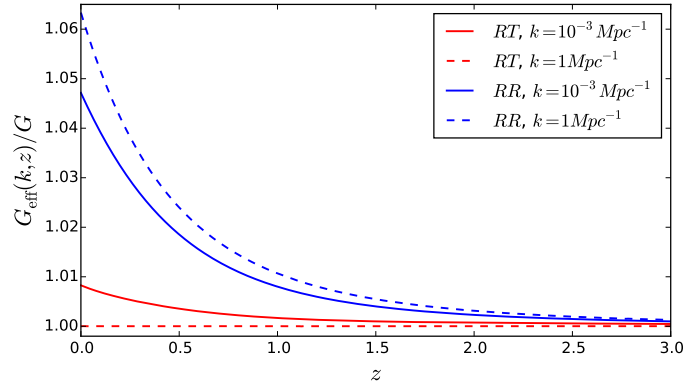


Figure 14: The effective Newton constant $G_{\text{eff}}(k, z)/G$.

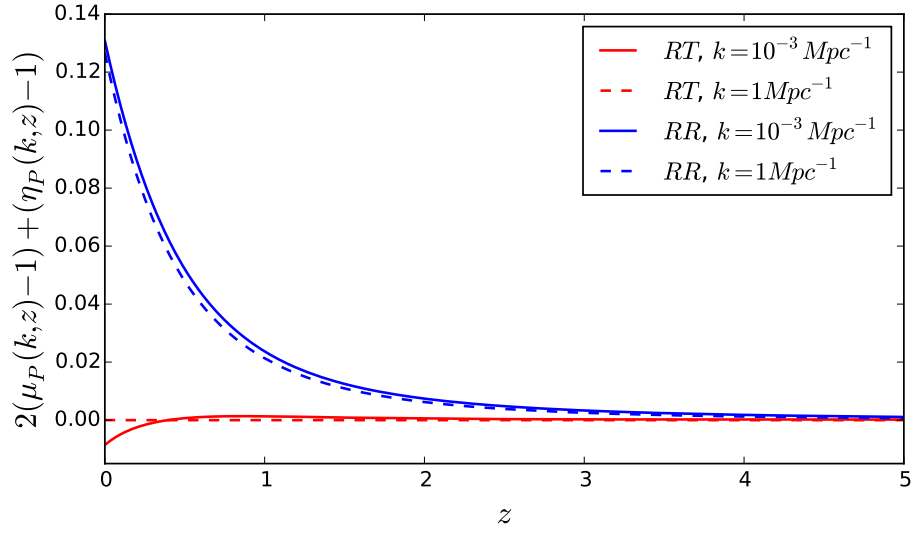


Figure 15: The combination used in Planck 2015 dark energy paper [84].

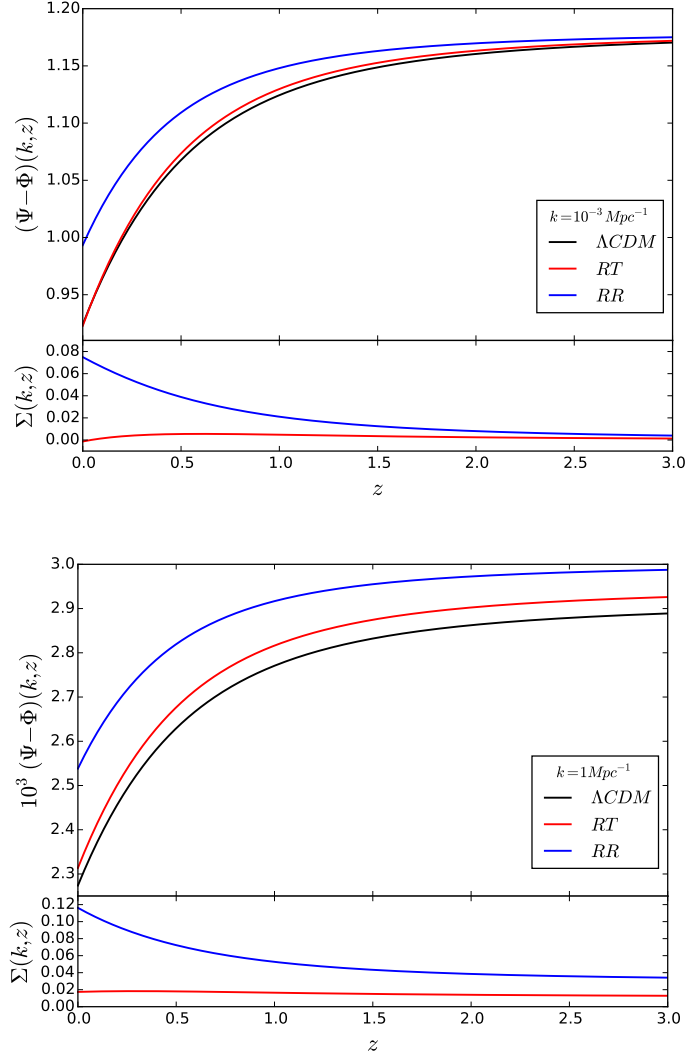


Figure 16: The difference of the potentials $\Psi - \Phi$ at large scales (upper) and at small scales (lower), and their relative ratio to ΛCDM .

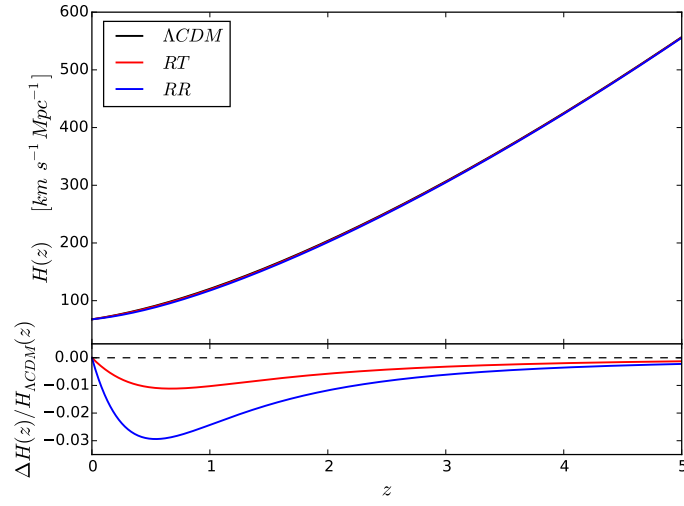


Figure 17: The Hubble parameter as a function of redshift z for the three models, using in all three cases the same values of the parameters, chosen to be the best-fit value for ΛCDM . In the upper panel the three curves are indistinguishable. In the lower panel we show the difference of the RT and RR models with respect to ΛCDM .

5 Tensor perturbations

Finally, we briefly discuss tensor perturbations in the RR and the RT model, to check for their stability. The stability of perturbations in the tensor sector is a priori non-trivial. We have mentioned for instance in the introduction that in bigravity a potentially interesting self-accelerating branch is in fact plagued by an instability in the tensor sector [31, 32]. For the RR and RT model, however, no such instability takes place.

For the RT model, in the tensor sector the equations governing the dynamics of the gravitational waves are given by

$$\partial_\eta^2 h_{ij}^{TT} + \left(2\mathcal{H} - 3\gamma\bar{V}H_0^2\right)\partial_\eta h_{ij}^{TT} + k^2 h_{ij}^{TT} = 8\pi G a^2 \Sigma_{ij}^{TT}. \quad (27)$$

Here \bar{V} is the background value of the auxiliary field V (see [66] for notation) and $\gamma = m^2/(9H_0^2)$, as in eq. (6), while the source term Σ_{ij}^{TT} is the traceless-transverse part of the anisotropic stress tensor. The latter is non-vanishing whenever relativistic free streaming particles are considered into the evolution [115, 116]. Decomposing the fields on the tensor eigenfunctions of the Laplacian, e.g. $h_{ij}^{TT} = H^{(T)} Q_{ij}^{(T)}$ [117], we can write,

$$\partial_\eta^2 H^{(T)} + \left(2\mathcal{H} - 3\gamma\bar{V}H_0^2\right)\partial_\eta H^{(T)} + k^2 H^{(T)} = 8\pi G a^2 \Sigma^{(T)}. \quad (28)$$

Similarly, for the RR model we get

$$\left(1 - 3\gamma\bar{V}\right)\left[\partial_\eta^2 H^{(T)} + 2\mathcal{H}\partial_\eta H^{(T)} + k^2 H^{(T)}\right] - 3\gamma\partial_\eta\bar{V}\partial_\eta H^{(T)} = 8\pi G a^2 \Sigma^{(T)}. \quad (29)$$

These equations have no unstable mode. Fig. 18 shows the result of the numerical integration of these equations, assuming an initial spectrum of perturbations corresponding to the one produced by single-field slow-roll inflation, where we indeed see that there is no instability in the tensor sector. These result have also been recently confirmed and extended in [59], where it has been found that, in contrast, the inclusion in the action of a term proportional to $C_{\mu\nu\rho\sigma}\Box^{-2}C^{\mu\nu\rho\sigma}$, where $C^{\mu\nu\rho\sigma}$ is the Weyl tensor, induces instabilities in the tensor sector.

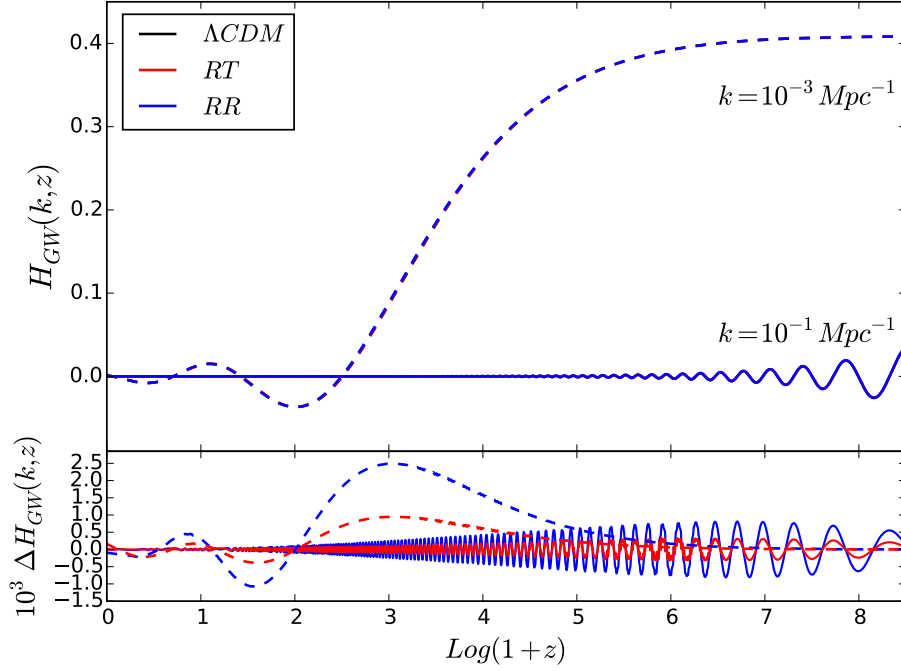


Figure 18: Upper panel: Amplitude of the gravitational waves (see (28), (29)) in ΛCDM (black), RT (red) and RR (blue) for $k = 10^{-1} \text{ Mpc}^{-1}$ (solid), $k = 10^{-3} \text{ Mpc}^{-1}$ (dashed), fixing $r = 0.05$. For each k , the three curves corresponding to the three models are indistinguishable from each other on the scale of this plot. Lower panel: Difference between the amplitude of the gravitational waves between RT and ΛCDM (red) and between RR and ΛCDM (blue) for $k = 10^{-1} \text{ Mpc}^{-1}$ (solid) and $k = 10^{-3} \text{ Mpc}^{-1}$ (dashed).

6 Conclusions

In this paper we have performed a detailed comparison of the RT and RR non-local models with cosmological data. We have implemented the cosmological perturbations of the model in a modified version of the CLASS Boltzmann code and we have compared these models to a rather complete set of data, including Planck 2015 data for temperature, polarization and lensing, isotropic and anisotropic BAO data, JLA supernovae, local measurement of H_0 , as well as structure formation data. The fact that these non-local models have reached the stage where such a detailed comparison is possible, is already by itself a rather non-trivial result. As we discussed in the introduction, from the purely phenomenological point of view no other attempt at modifying gravity in the infrared has reached this stage. We have found that, by themselves, both the RR and RT non-local models provide good fits to the data. We have then compared the performances of the two non-local models between them and with ΛCDM . The result is that, with present data, ΛCDM and the RT model are statistically indistinguishable. They both fit the data extremely well. This is remarkable, considering that they are genuinely different models. The non-local models are not extensions of ΛCDM with extra free parameters. Rather, they have the same

number of parameters as Λ CDM, with a mass scale replacing the cosmological constant, and there is no formal limit in which they reduce to Λ CDM. The differences between the predictions of the RT model and those of Λ CDM are of order of a few percent, for instance in structure formation; this is a level which allows both models to be consistent with present data, although they could in principle be distinguishable with the data of future missions. The RR model, in contrast, performs less well in the comparison with data, compared both to Λ CDM and to the RT model.

We should however stress that the comparison between the models generically depends on the prior assumed and in particular on the values of the sum of the neutrino masses, which could potentially be relevant in our case. In this paper we have used the sum of the neutrino masses of the *Planck* baseline analysis [83], i.e. a normal mass hierarchy with $\sum_\nu m_\nu = 0.06$ eV. We make this choice in order to facilitate the comparison with the *Planck* analysis of Λ CDM. However, oscillation experiments can be consistent with larger neutrino masses, including an inverted hierarchy with $\sum_\nu m_\nu \gtrsim 0.1$ eV. The effect of increasing the neutrino masses is to lower H_0 , which in Λ CDM would increase the tension with local measurements. However, we have seen that the non-local models, and particularly the RR model, predict higher values of H_0 , so they can accommodate larger neutrino masses, without entering in tension with local measurements. Furthermore, larger neutrino masses also have the effect of lowering σ_8 (see [83] and refs. therein). In view of the result shown in Fig. 10, this would improve the agreement of the non-local models with structure formation data. In particular one can expect that, leaving the neutrino masses as free parameters, within the priors allowed by oscillation experiments, should improve the agreement with the data of the RR model, rendering it again competitive with Λ CDM. This possibility is currently being investigated in [98].

These results encourage further studies of this class of non-local models, and give precious hints for the search of the mechanism that generates the non-localities from a fundamental local theory.

Acknowledgments. We thank Benjamin Audren, Alex Barreira, Julien Lesgourgues, Baojiu Li, and Ignacy Sawiki for useful discussions. The work of YD, SF, MK and MM is supported by the Fonds National Suisse. VP acknowledges the DFG TransRegio TRR33 grant on The Dark Universe. The work of MM is supported by the SwissMap NCCR. Part of the numerical work presented in this publication used the Baobab cluster of the University of Geneva.

The development of Planck has been supported by: ESA; CNES and CNRS/INSU-IN2P3-INP (France); ASI, CNR, and INAF (Italy); NASA and DoE (USA); STFC and UKSA (UK); CSIC, MICINN and JA (Spain); Tekes, AoF and CSC (Finland); DLR and MPG (Germany); CSA (Canada); DTU Space (Denmark); SER/SSO (Switzerland); RCN (Norway); SFI (Ireland); FCT/MCTES (Portugal). A description of the Planck Collaboration and a list of its members, including the technical or scientific activities in which they have been involved, can be found at <http://www.cosmos.esa.int/web/planck/planck-collaboration>.

A Implementation of the nonlocal models in CLASS

The full set of cosmological equations of motion for the RT and RR models, at the background and scalar linear perturbations level, has already been given in [66]. In this appendix we briefly comment on the global strategy used in CLASS for evolving them. For completeness we also review here the equations used in the code, in the corresponding format. The code itself will become publicly available soon [80].

CLASS uses the following convention for scalar perturbations¹³

$$ds^2 = a^2 \left[- (1 + 2\psi) d\eta^2 + [(1 - 2\phi)\delta_{ij}] dx^i dx^j \right], \quad (30)$$

which is different from the one we adopted in eq. (21). In order to facilitate the comparison with the code, we will use CLASS convention in this appendix. Since the nonlocal models both have the same overall structure, that is, a set of modified Einstein equations including auxiliary fields along with second order differential equations dictating the evolution of the latter, their implementation in CLASS is similar.

A.1 Implementation of the RT model

We start by writing down the relevant cosmological background equations corresponding to the RT model [52, 65]. In the code, the (00) component of Einstein equations is used to infer algebraically the Hubble parameter $H \equiv a'/a^2$ in terms of the energy density of the matter component that one wishes to take into account. In our case it reads,

$$H = \frac{\gamma \bar{V} H_0^2}{2a} + \left[\left(\frac{\gamma \bar{V} H_0^2}{2a} \right) + \gamma (\bar{U} - \bar{V}'/a^2) H_0^2 + \frac{8\pi G}{3} \bar{\rho} \right]^{1/2}, \quad (31)$$

where H_0 is the present value of the Hubble parameter and \bar{U} and \bar{V} denote the background values of the auxiliary fields U and V . The derivative of the Hubble parameter is not computed numerically, but rather obtained algebraically from the trace of the (ij) component of Einstein equations, in terms of the pressure,

$$H' = -\frac{3}{2}a \left[H^2 + \frac{8\pi G}{3} \bar{p} + \gamma \left(\frac{H}{a} \bar{V} - \bar{U} \right) H_0^2 \right]. \quad (32)$$

As already mentioned above, in our case, the evolution of the auxiliary fields is dictated by a set of second order differential equations,

$$\bar{U}'' + 2aH\bar{U}' = 6a(H' + 2aH^2) \quad , \quad \bar{V}'' - a(H' + 5aH^2)\bar{V} = a^2\bar{U}' \quad (33)$$

which need to be integrated numerically directly within the code with initial conditions $\bar{U} = \bar{U}' = \bar{V} = \bar{V}' = 0$ set deep into the radiation era [52, 65]. Once the matter components have been specified, such that their energy density and pressure can be written explicitly (for example by using their separate energy-momentum conservation together with an equation of state, or their unperturbed phase-space distribution functions) the system closes and the background evolution can be integrated.

¹³At the time of writing the present paper, the equations for the nonlocal models have been implemented only in Newtonian gauge into CLASS, and not in the synchronous one.

Concerning the linear scalar perturbations, the set of equations needed includes the conservation equation of each fluid component, along with the evolution equations for the auxiliary fields perturbations and two independent components of the modified Einstein equations. For the latter, the code originally uses the divergence of the $(0i)$ component in order to extract ϕ' algebraically. In our case we have¹⁴

$$\phi' = -\mathcal{H}\psi + 4\pi G \frac{a^2}{k^2}(\bar{\rho} + \bar{p})\theta + \frac{3}{2}\gamma \left[\mathcal{H}\delta Z - \frac{1}{2}\delta Z' + \frac{1}{2}\psi\bar{V} - \frac{1}{2}\delta V \right] H_0^2. \quad (34)$$

The expression of ψ is obtained in terms of ϕ from the longitudinal-traceless part of the (ij) component of Einstein equation (the one sourced by the scalar anisotropic stress σ , see eq. (4.8)-(4.9) of [66] for more details),

$$\psi = \phi - 12\pi G \frac{a^2}{k^2}(\bar{\rho} + \bar{p})\sigma + 3\gamma\delta Z H_0^2, \quad (35)$$

and ϕ is obtained from (34) by numerical integration. The equations governing the dynamics of the auxiliary fields are,

$$\delta U'' + k^2\delta U + 2\mathcal{H}\delta U' = (\psi' + 3\phi')(\bar{U}' - 6\mathcal{H}) + 2k^2(\psi - 2\phi) - 6\phi'', \quad (36)$$

$$\delta Z'' - 2(\mathcal{H}' + 2\mathcal{H}^2 - k^2)\delta Z = 2a^2\delta U - 4\mathcal{H}\delta V - \delta V' + 3\bar{V}'\psi + (\psi' + 2\phi')\bar{V}, \quad (37)$$

$$\delta V'' - \left(\mathcal{H}' + 4\mathcal{H}^2 - \frac{k^2}{2} \right) \delta V = \quad (38)$$

$$a^2\delta U' - \frac{1}{2}k^2\delta Z' + 2\mathcal{H}k^2\delta Z + \left[\frac{k^2}{2}\psi - \mathcal{H}(\psi' + 9\phi') \right] \bar{V} + (\psi' + 3\phi')\bar{V}' + a^2\bar{U}'\psi.$$

We see that they contain ψ' and ϕ'' , which can also be extracted algebraically from the equations. In particular, ϕ'' is obtained from the trace of the (ij) component of Einstein equations,

$$\begin{aligned} \phi'' = & -\psi(\mathcal{H}^2 + 2\mathcal{H}') - \mathcal{H}(\psi' + 2\phi') + \frac{k^2}{3}(\psi - \phi) \\ & - \frac{3}{2} \left[\gamma \left(a^2\delta U + (\phi' + \mathcal{H}\psi)\bar{V} - \mathcal{H}\delta V - \frac{k^2}{3}\delta Z \right) H_0^2 - \frac{8\pi G}{3}a^2\delta p \right], \end{aligned} \quad (39)$$

whereas ψ' is obtained by taking the derivatives of (35),

$$\psi' = \phi' - \frac{24\pi G}{k^2}\mathcal{H}a^2(\bar{\rho} + \bar{p})\sigma - 12\pi G \frac{a^2}{k^2}[(\bar{\rho} + \bar{p})\sigma]' + 3\gamma\delta Z' H_0^2. \quad (40)$$

Within the version of CLASS that we used, the source term $[(\bar{\rho} + \bar{p})\sigma]$ decomposes as

$$[(\bar{\rho} + \bar{p})\sigma] = \frac{4}{3}(\rho_\gamma\sigma_\gamma + \rho_{ur}\sigma_{ur}) + (\rho + p)_{ncdm}\sigma_{ncdm}. \quad (41)$$

¹⁴Here $\mathcal{H} \equiv a'/a$, as used in the perturbation module of CLASS.

where γ , ur and $ncdm$ denote the photon, ultra-relativistic particles and non-cold dark matter (NCDM) components respectively. Therefore to compute its corresponding time derivative one needs,

$$[(\bar{\rho} + \bar{p})\sigma]' = \frac{4}{3}(\rho'_\gamma\sigma_\gamma + \rho'_{ur}\sigma_{ur} + \rho_\gamma\sigma'_\gamma + \rho_{ur}\sigma'_{ur}) + [(\rho + p)_{ncdm}\sigma_{ncdm}]', \quad (42)$$

and, since at background level each of the ultra-relativistic particles and photons components is conserved, one can use the background conservation equation, $\bar{\rho}' = -3\mathcal{H}(\bar{\rho} + \bar{p})$ together with the equation of state for ultra-relativistic particle $\bar{p} = (1/3)\bar{\rho}$, to write,

$$[(\bar{\rho} + \bar{p})\sigma]' = \frac{4}{3}\rho_\gamma(\sigma'_\gamma - 4\mathcal{H}\sigma_\gamma) + \frac{4}{3}\rho_{ur}(\sigma'_{ur} - 4\mathcal{H}\sigma_{ur}) + [(\rho + p)_{ncdm}\sigma_{ncdm}]', \quad (43)$$

while the (massive) NCDM components have to be treated separately in terms of their phase-space description. This form is quite convenient for being implemented into the code since the various quantities have already been built into the original version.

A.2 Implementation of the RR model

In the case of the RR model the general structure is the same as above and we only display the relevant equations needed for the implementation in the code. At the background level, the (00) component of the modified Einstein equations [58] leads to

$$H = \left(1 - 3\gamma\bar{V}\right)^{-1} \left\{ \frac{3\gamma}{2a}\bar{V}' + \left[\left(\frac{3\gamma}{2a}\bar{V}' \right)^2 + \left(1 - 3\gamma\bar{V}\right) \left(\frac{\gamma}{4}\bar{U}^2 H_0^2 - \frac{\gamma}{2a^2}\bar{V}'\bar{U}' + \frac{8\pi G}{3}\rho \right) \right]^{1/2} \right\}. \quad (44)$$

From the (ij) component one gets

$$H' = -\frac{3}{2}a \left\{ H^2 + \left(1 - 3\gamma\bar{V}\right)^{-1} \left[\frac{8\pi G}{3}\bar{p} - \gamma \left(\frac{1}{4}\bar{U}^2 H_0^2 + \bar{U} H_0^2 - \frac{H}{a}\bar{V}' + \frac{1}{2a^2}\bar{V}'\bar{U}' \right) \right] \right\}. \quad (45)$$

The equation of the auxiliary fields are

$$\bar{U}'' + 2aH\bar{U}' = 6a(H' + 2aH^2), \quad \bar{V}'' + 2aH\bar{V}' = a^2\bar{U}H_0^2. \quad (46)$$

To linear order in the scalar perturbations, the (0*i*) component of the modified Einstein equations leads to,

$$\phi' = -\mathcal{H}\psi + \frac{3}{2}\left(1 - 3\gamma\bar{V}\right)^{-1} \left[\frac{8\pi Ga^2}{3k^2}(\bar{\rho} + \bar{p})\theta - \gamma \left(\delta V' - \bar{V}'\psi - \mathcal{H}\delta V + \frac{1}{2}(\bar{U}'\delta V + \bar{V}'\delta U) \right) \right]. \quad (47)$$

The equation sourced by the anisotropic stress gives

$$\psi = \phi + \left(1 - 3\gamma\bar{V}\right)^{-1} \left[-12\pi G \frac{a^2}{k^2}(\bar{\rho} + \bar{p})\sigma + 3\gamma\delta V \right]. \quad (48)$$

The perturbation of the equations for the auxiliary field equations gives

$$\delta U'' + 2\mathcal{H}\delta U' + k^2\delta U = (\psi' + 3\phi')(\bar{U}' - 6\mathcal{H}) + 2k^2(\psi - 2\phi) - 6\phi'', \quad (49)$$

$$\delta V'' + 2\mathcal{H}\delta V' + k^2\delta V = (\psi' + 3\phi')\bar{V}' + 2a^2\psi\bar{U}H_0^2 + a^2\delta UH_0^2. \quad (50)$$

In order to solve the system, one has also to provide expressions of ψ' and ϕ'' obtained similarly to the case of the RT model. One finds

$$\psi' = \phi' + \left(1 - 3\gamma\bar{V}\right)^{-1} \left[3\gamma(\psi - \phi)\bar{V}' - \frac{24\pi G}{k^2}\mathcal{H}a^2(\bar{\rho} + \bar{p})\sigma - 12\pi G\frac{a^2}{k^2}[(\bar{\rho} + \bar{p})\sigma]' + 3\gamma\delta V' \right], \quad (51)$$

and

$$\begin{aligned} \phi'' = & -\psi(\mathcal{H}^2 + 2\mathcal{H}') - \mathcal{H}(\psi' + 2\phi') + \frac{k^2}{3}(\psi - \phi) \\ & - \frac{3}{2}\left(1 - 3\gamma\bar{V}\right)^{-1} \left\{ \gamma \left[\frac{1}{2}a^2\bar{U}\delta UH_0^2 - 2a^2\psi\bar{U}H_0^2 - (2\phi' - 2\mathcal{H}\psi + \psi' + \psi\bar{U}')\bar{V}' \right. \right. \\ & \left. \left. + \delta V'' + \mathcal{H}\delta V' + \left(\mathcal{H}^2 + 2\mathcal{H}' + \frac{2k^2}{3} \right) \delta V + \frac{1}{2}(\bar{U}'\delta V' + \bar{V}'\delta U') \right] - \frac{8\pi G}{3}a^2\delta p \right\}. \end{aligned} \quad (52)$$

B Comparison with Lunar Laser Ranging

Ref. [64] has raised the question of the consistency of the RR model with Lunar Laser Ranging experiments. The concern arises from the fact that, in the RR model, among the equations obtained perturbing over the cosmological solution, we have a Poisson equation where the term $\nabla^2\Phi$ is multiplied by a factor $[1 - (m^2\bar{S}/3)]$, where \bar{S} is the background value of the auxiliary field S (see [66] for definitions and notation), leading to an effective Newton constant

$$G_{\text{eff}} = G \left[1 - \frac{m^2\bar{S}}{3} \right]^{-1} \left[1 + \mathcal{O}\left(\frac{H_0^2}{k^2}\right) \right], \quad (53)$$

see eqs. (23,24) of [64] or eq. (5.3) of [66]. If one now uses the time dependence of $\bar{S}(t)$ found from the cosmological background solution, and applies eq. (53) to the Earth-Moon system, one finds a value $\dot{G}_{\text{eff}}/G \simeq 92 \times 10^{-13} \text{ yr}^{-1}$ [64], which exceeds the current bound from Lunar Laser Ranging (LLR), $\dot{G}_{\text{eff}}/G \simeq (4 \pm 9) \times 10^{-13} \text{ yr}^{-1}$ [118].

As a first remark, we observe that this potential problem does not appear in the RT model. Indeed, for the RT model inside the horizon $G_{\text{eff}}/G = 1 + \mathcal{O}(H_0^2/k^2)$, see [63,66,67], so for the RT model the issue raised in [64] does not apply, and there is no problem with LLR.

For the RR model the issue is more delicate. An obvious question, indeed mentioned in [64], is whether one is allowed to use the background expansion rate for the field $\bar{S}(t)$, computed at cosmological scales, inside a system such as the Earth-Moon system. The Earth-Moon system, and more generally the solar system or even the Local Group, are not expanding with the Hubble flow, so the numerical estimate suggesting a value $\dot{G}_{\text{eff}}/G \simeq 92 \times 10^{-13} \text{ yr}^{-1}$ cannot be taken literally. Nevertheless, there still remains the possibility

that, even if the background geometry at local scales (i.e. at the scales of the Earth-Moon system, or of the solar system) no longer follows the cosmological expansion, a scalar field could still have a residual time dependence inherited from its behavior at cosmological scales. Indeed, this can happen in scalar-tensor theories in which the scalar field φ has a shift symmetry $\varphi \rightarrow \varphi + \text{const.}$ [119]. In this case the equations of motion admit a separation of variables of the form $\varphi(t, r) = \varphi(r) + \varphi_{\text{cosm}}(t)$, with $\dot{\varphi}_{\text{cosm}}(t) \sim H(t)$. Since the LLR bound, expressed in terms of H_0 , reads $\dot{G}/G = (0.6 \pm 1.3) \times 10^{-2} H_0$, if this scalar field enters in the effective Newton constant (as happens in most tensor-scalar theories, as well as in the RR model, but not in the RT model), this could induce, depending on the precise numerical factor, a violation of the LLR bound.¹⁵

In our case, in the RR model there is no shift symmetry $S \rightarrow S + \text{const.}$ for the auxiliary field S (see e.g. eqs. (4)-(7) of [58]), so the above argument does not apply directly. Nevertheless, for the RR model, in the absence of an explicit study of the equations including as sources both a cosmological fluid and a localized static source, one cannot exclude a priori the presence of a residual time dependence in the auxiliary field S , on local scales. Similarly, it would be interesting to study the stability of such r - and t -dependent solution under time-dependent perturbations. We leave these issues for future work.

References

- [1] P. Bull *et. al.*, “Beyond Λ CDM: Problems, solutions, and the road ahead,” [1512.05356](#).
- [2] G. Dvali, G. Gabadadze, and M. Porrati, “4-D gravity on a brane in 5-D Minkowski space,” *Phys.Lett.* **B485** (2000) 208–214, [hep-th/0005016](#).
- [3] C. de Rham and G. Gabadadze, “Generalization of the Fierz-Pauli Action,” *Phys.Rev.* **D82** (2010) 044020, [1007.0443](#).
- [4] C. de Rham, G. Gabadadze, and A. J. Tolley, “Resummation of Massive Gravity,” *Phys.Rev.Lett.* **106** (2011) 231101, [1011.1232](#).
- [5] S. Hassan and R. A. Rosen, “Resolving the Ghost Problem in non-Linear Massive Gravity,” *Phys.Rev.Lett.* **108** (2012) 041101, [1106.3344](#).
- [6] K. Hinterbichler, “Theoretical Aspects of Massive Gravity,” *Rev.Mod.Phys.* **84** (2012) 671–710, [1105.3735](#).
- [7] C. de Rham, “Massive Gravity,” *Living Rev. Rel.* **17** (2014) 7, [1401.4173](#).
- [8] S. Hassan and R. A. Rosen, “Bimetric Gravity from Ghost-free Massive Gravity,” *JHEP* **1202** (2012) 126, [1109.3515](#).
- [9] A. Schmidt-May and M. von Strauss, “Recent developments in bimetric theory,” [1512.00021](#).

¹⁵Note however that it is possible that this ansatz leads to an differential equation for $\varphi(r)$ which has no regular solution [119], in which case this would not be the solution relevant to the physics of the problem, i.e. the solution corresponding to a metric interpolating from a Schwarzschild-like behavior in the local region to the FRW behavior at cosmological distances.

- [10] C. Deffayet, “Cosmology on a brane in Minkowski bulk,” *Phys.Lett.* **B502** (2001) 199–208, [hep-th/0010186](#).
- [11] C. Deffayet, G. Dvali, and G. Gabadadze, “Accelerated universe from gravity leaking to extra dimensions,” *Phys.Rev.* **D65** (2002) 044023, [astro-ph/0105068](#).
- [12] M. A. Luty, M. Porrati, and R. Rattazzi, “Strong interactions and stability in the DGP model,” *JHEP* **0309** (2003) 029, [hep-th/0303116](#).
- [13] A. Nicolis and R. Rattazzi, “Classical and quantum consistency of the DGP model,” *JHEP* **0406** (2004) 059, [hep-th/0404159](#).
- [14] D. Gorbunov, K. Koyama, and S. Sibiryakov, “More on ghosts in DGP model,” *Phys.Rev.* **D73** (2006) 044016, [hep-th/0512097](#).
- [15] C. Charmousis, R. Gregory, N. Kaloper, and A. Padilla, “DGP Spectroscopy,” *JHEP* **0610** (2006) 066, [hep-th/0604086](#).
- [16] K. Izumi, K. Koyama, and T. Tanaka, “Unexorcized ghost in DGP brane world,” *JHEP* **0704** (2007) 053, [hep-th/0610282](#).
- [17] G. D’Amico, C. de Rham, S. Dubovsky, G. Gabadadze, D. Pirtskhalava, and A. J. Tolley, “Massive Cosmologies,” *Phys.Rev.* **D84** (2011) 124046, [1108.5231](#).
- [18] A. De Felice, A. E. Gumrukcuoglu, C. Lin, and S. Mukohyama, “Nonlinear stability of cosmological solutions in massive gravity,” *JCAP* **1305** (2013) 035, [1303.4154](#).
- [19] M. S. Volkov, “Cosmological solutions with massive gravitons in the bigravity theory,” *JHEP* **01** (2012) 035, [1110.6153](#).
- [20] D. Comelli, M. Crisostomi, F. Nesti, and L. Pilo, “FRW Cosmology in Ghost Free Massive Gravity,” *JHEP* **03** (2012) 067, [1111.1983](#). [Erratum: JHEP06,020(2012)].
- [21] M. von Strauss, A. Schmidt-May, J. Enander, E. Mortsell, and S. Hassan, “Cosmological Solutions in Bimetric Gravity and their Observational Tests,” *JCAP* **1203** (2012) 042, [1111.1655](#).
- [22] Y. Akrami, T. S. Koivisto, and M. Sandstad, “Accelerated expansion from ghost-free bigravity: a statistical analysis with improved generality,” *JHEP* **03** (2013) 099, [1209.0457](#).
- [23] N. Tamanini, E. N. Saridakis, and T. S. Koivisto, “The Cosmology of Interacting Spin-2 Fields,” *JCAP* **1402** (2014) 015, [1307.5984](#).
- [24] M. Fasiello and A. J. Tolley, “Cosmological Stability Bound in Massive Gravity and Bigravity,” *JCAP* **1312** (2013) 002, [1308.1647](#).
- [25] Y. Akrami, T. S. Koivisto, D. F. Mota, and M. Sandstad, “Bimetric gravity doubly coupled to matter: theory and cosmological implications,” *JCAP* **1310** (2013) 046, [1306.0004](#).

- [26] F. Koennig, A. Patil, and L. Amendola, “Viable cosmological solutions in massive bimetric gravity,” *JCAP* **1403** (2014) 029, [1312.3208](#).
- [27] D. Comelli, M. Crisostomi, and L. Pilo, “FRW Cosmological Perturbations in Massive Bigravity,” [1403.5679](#).
- [28] A. R. Solomon, Y. Akrami, and T. S. Koivisto, “Cosmological perturbations in massive bigravity: I. Linear growth of structures,” [1404.4061](#).
- [29] A. De Felice, A. E. Gumrukcuoglu, S. Mukohyama, N. Tanahashi, and T. Tanaka, “Viable cosmology in bimetric theory,” *JCAP* **1406** (2014) 037, [1404.0008](#).
- [30] F. Koennig and L. Amendola, “A minimal bimetric gravity model that fits cosmological observations,” [1402.1988](#).
- [31] M. Lagos and P. G. Ferreira, “Cosmological perturbations in massive bigravity,” *JCAP* **1412** (2014) 026, [1410.0207](#).
- [32] G. Cusin, R. Durrer, P. Guarato, and M. Motta, “Gravitational waves in bigravity cosmology,” *JCAP* **1505** (2015) 030, [1412.5979](#).
- [33] F. Koennig, “Higuchi Ghosts and Gradient Instabilities in Bimetric Gravity,” *Phys. Rev.* **D91** (2015) 104019, [1503.07436](#).
- [34] D. Comelli, M. Crisostomi, and L. Pilo, “Perturbations in Massive Gravity Cosmology,” *JHEP* **1206** (2012) 085, [1202.1986](#).
- [35] F. Koennig, Y. Akrami, L. Amendola, M. Motta, and A. R. Solomon, “Stable and unstable cosmological models in bimetric massive gravity,” *Phys. Rev.* **D90** (2014) 124014, [1407.4331](#).
- [36] Y. Akrami, S. F. Hassan, F. Könnig, A. Schmidt-May, and A. R. Solomon, “Bimetric gravity is cosmologically viable,” *Phys. Lett.* **B748** (2015) 37–44, [1503.07521](#).
- [37] S. Deser and R. Woodard, “Nonlocal Cosmology,” *Phys.Rev.Lett.* **99** (2007) 111301, [0706.2151](#).
- [38] S. Deser and R. Woodard, “Observational Viability and Stability of Nonlocal Cosmology,” *JCAP* **1311** (2013) 036, [1307.6639](#).
- [39] R. Woodard, “Nonlocal Models of Cosmic Acceleration,” *Found.Phys.* **44** (2014) 213–233, [1401.0254](#).
- [40] A. Barvinsky, “Nonlocal action for long distance modifications of gravity theory,” *Phys.Lett.* **B572** (2003) 109, [hep-th/0304229](#).
- [41] A. Barvinsky, “Dark energy and dark matter from nonlocal ghost-free gravity theory,” *Phys.Lett.* **B710** (2012) 12–16, [1107.1463](#).
- [42] A. O. Barvinsky, “Serendipitous discoveries in nonlocal gravity theory,” *Phys.Rev.* **D85** (2012) 104018, [1112.4340](#).

- [43] T. Koivisto, “Dynamics of Nonlocal Cosmology,” *Phys.Rev.* **D77** (2008) 123513, [0803.3399](#).
- [44] T. Koivisto, “Newtonian limit of nonlocal cosmology,” *Phys.Rev.* **D78** (2008) 123505, [0807.3778](#).
- [45] S. Capozziello, E. Elizalde, S. Nojiri, and S. D. Odintsov, “Accelerating cosmologies from non-local higher-derivative gravity,” *Phys.Lett.* **B671** (2009) 193–198, [0809.1535](#).
- [46] E. Elizalde, E. Pozdeeva, and S. Y. Vernov, “De Sitter Universe in Non-local Gravity,” *Phys.Rev.* **D85** (2012) 044002, [1110.5806](#).
- [47] Y. Zhang and M. Sasaki, “Screening of cosmological constant in non-local cosmology,” *Int.J.Mod.Phys.* **D21** (2012) 1250006, [1108.2112](#).
- [48] E. Elizalde, E. Pozdeeva, and S. Y. Vernov, “Reconstruction Procedure in Nonlocal Models,” *Class.Quant.Grav.* **30** (2013) 035002, [1209.5957](#).
- [49] S. Park and S. Dodelson, “Structure formation in a nonlocally modified gravity model,” *Phys.Rev.* **D87** (2013) 024003, [1209.0836](#).
- [50] K. Bamba, S. Nojiri, S. D. Odintsov, and M. Sasaki, “Screening of cosmological constant for De Sitter Universe in non-local gravity, phantom-divide crossing and finite-time future singularities,” *Gen.Rel.Grav.* **44** (2012) 1321–1356, [1104.2692](#).
- [51] S. Dodelson and S. Park, “Nonlocal Gravity and Structure in the Universe,” *Phys.Rev.* **D90** (2014) 043535, [1310.4329](#).
- [52] M. Maggiore, “Phantom dark energy from nonlocal infrared modifications of general relativity,” *Phys.Rev.* **D89** (2014) 043008, [1307.3898](#).
- [53] N. Arkani-Hamed, S. Dimopoulos, G. Dvali, and G. Gabadadze, “Nonlocal modification of gravity and the cosmological constant problem,” [hep-th/0209227](#).
- [54] G. Dvali, “Predictive Power of Strong Coupling in Theories with Large Distance Modified Gravity,” *New J.Phys.* **8** (2006) 326, [hep-th/0610013](#).
- [55] G. Dvali, S. Hofmann, and J. Khoury, “Degravitation of the cosmological constant and graviton width,” *Phys.Rev.* **D76** (2007) 084006, [hep-th/0703027](#).
- [56] M. Porrati, “Fully covariant van Dam-Veltman-Zakharov discontinuity, and absence thereof,” *Phys.Lett.* **B534** (2002) 209, [hep-th/0203014](#).
- [57] M. Jaccard, M. Maggiore, and E. Mitsou, “A non-local theory of massive gravity,” *Phys.Rev.* **D88** (2013) 044033, [1305.3034](#).
- [58] M. Maggiore and M. Mancarella, “Non-local gravity and dark energy,” *Phys.Rev.* **D90** (2014) 023005, [1402.0448](#).
- [59] G. Cusin, S. Foffa, M. Maggiore, and M. Mancarella, “Nonlocal gravity with a Weyl-square term,” *Phys. Rev.* **D93** (2016) 043006, [1512.06373](#).

- [60] P. G. Ferreira and A. L. Maroto, “A few cosmological implications of tensor nonlocalities,” *Phys.Rev.* **D88** (2013) 123502, [1310.1238](#).
- [61] M. Maggiore, “Dark energy and dimensional transmutation in R^2 gravity,” [1506.06217](#).
- [62] M. Maggiore, “Perturbative loop corrections and nonlocal gravity,” *Phys. Rev.* **D93** (2016) 063008, [1603.01515](#).
- [63] A. Kehagias and M. Maggiore, “Spherically symmetric static solutions in a non-local infrared modification of General Relativity,” *JHEP* **1408** (2014) 029, [1401.8289](#).
- [64] A. Barreira, B. Li, W. A. Hellwing, C. M. Baugh, and S. Pascoli, “Nonlinear structure formation in Nonlocal Gravity,” *JCAP* **1409** (2014) 031, [1408.1084](#).
- [65] S. Foffa, M. Maggiore, and E. Mitsou, “Cosmological dynamics and dark energy from non-local infrared modifications of gravity,” *Int.J.Mod.Phys.* **A29** (2014) 1450116, [1311.3435](#).
- [66] Y. Dirian, S. Foffa, N. Khosravi, M. Kunz, and M. Maggiore, “Cosmological perturbations and structure formation in nonlocal infrared modifications of general relativity,” *JCAP* **1406** (2014) 033, [1403.6068](#).
- [67] S. Nesseris and S. Tsujikawa, “Cosmological perturbations and observational constraints on nonlocal massive gravity,” *Phys.Rev.* **D90** (2014) 024070, [1402.4613](#).
- [68] L. Modesto and S. Tsujikawa, “Non-local massive gravity,” *Phys.Lett.* **B727** (2013) 48–56, [1307.6968](#).
- [69] S. Foffa, M. Maggiore, and E. Mitsou, “Apparent ghosts and spurious degrees of freedom in non-local theories,” *Phys.Lett.* **B733** (2014) 76–83, [1311.3421](#).
- [70] A. Conroy, T. Koivisto, A. Mazumdar, and A. Teimouri, “Generalized quadratic curvature, non-local infrared modifications of gravity and Newtonian potentials,” *Class. Quant. Grav.* **32** (2015) 015024, [1406.4998](#).
- [71] G. Cusin, J. Fumagalli, and M. Maggiore, “Non-local formulation of ghost-free bigravity theory,” *JHEP* **1409** (2014) 181, [1407.5580](#).
- [72] Y. Dirian and E. Mitsou, “Stability analysis and future singularity of the $m^2 R \square^{-2} R$ model of non-local gravity,” *JCAP* **10** (2014) 065, [1408.5058](#).
- [73] E. Mitsou, *Aspects of Infrared Non-local Modifications of General Relativity*. PhD thesis, Geneva U., 2015. [1504.04050](#).
- [74] A. Barreira, B. Li, E. Jennings, J. Merten, L. King, *et. al.*, “Galaxy cluster lensing masses in modified lensing potentials,” *Mon. Not. Roy. Astron. Soc.* **454** (2015) 4085, [1505.03468](#).

- [75] A. Barreira, M. Cautun, B. Li, C. Baugh, and S. Pascoli, “Weak lensing by voids in modified lensing potentials,” *JCAP* **1508** (2015) 028, [1505.05809](#).
- [76] A. Codello and R. K. Jain, “Covariant Effective Field Theory of Gravity II: Cosmological Implications,” [1507.07829](#).
- [77] G. Cusin, S. Foffa, M. Maggiore, and M. Mancarella, “Conformal symmetry and nonlinear extensions of nonlocal gravity,” *Phys. Rev.* **D93** (2016) 083008, [1602.01078](#).
- [78] A. Codello and R. K. Jain, “A Unified Evolution of the Universe,” [1603.00028](#).
- [79] Y. Dirian, S. Foffa, M. Kunz, M. Maggiore, and V. Pettorino, “Non-local gravity and comparison with observational datasets,” *JCAP* **1504** (2015) 044, [1411.7692](#).
- [80] https://github.com/dirian/class_public/tree/nonlocal.
- [81] D. Blas, J. Lesgourgues, and T. Tram, “The Cosmic Linear Anisotropy Solving System (CLASS). Part II: Approximation schemes,” *JCAP* **7** (July, 2011) 34, [1104.2933](#).
- [82] B. Audren, J. Lesgourgues, K. Benabed, and S. Prunet, “Conservative constraints on early cosmology with MONTE PYTHON,” *JCAP* **2** (Feb., 2013) 1, [1210.7183](#).
- [83] **Planck** Collaboration, P. Ade *et. al.*, “Planck 2015 results. XIII. Cosmological parameters,” [1502.01589](#).
- [84] **Planck** Collaboration, P. A. R. Ade *et. al.*, “Planck 2015 results. XIV. Dark energy and modified gravity,” [1502.01590](#).
- [85] **Planck** Collaboration, R. Adam *et. al.*, “Planck 2015 results. I. Overview of products and scientific results,” [1502.01582](#).
- [86] **Planck** Collaboration, P. A. R. Ade *et. al.*, “Planck 2013 results. I. Overview of products and scientific results,” *Astron. Astrophys.* **571** (2014) A1, [1303.5062](#).
- [87] **Planck** Collaboration, “Planck 2015 release explanatory supplement,” http://wiki.cosmos.esa.int/planckpla2015/index.php/CMB_spectrum%26Likelihood_Code.
- [88] **Planck** Collaboration, N. Aghanim *et. al.*, “Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters,” *Submitted to: Astron. Astrophys.* (2015) [1507.02704](#).
- [89] **Planck** Collaboration, P. A. R. Ade *et. al.*, “Planck 2015 results. XV. Gravitational lensing,” [1502.01591](#).
- [90] **SDSS** Collaboration, M. Betoule *et. al.*, “Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples,” *Astron. Astrophys.* **568** (2014) A22, [1401.4064](#).

- [91] F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, L. Campbell, Q. Parker, W. Saunders, and F. Watson, “The 6dF Galaxy Survey: baryon acoustic oscillations and the local Hubble constant,” *Mon.Not.Roy.Astron.Soc.* **416** (Oct., 2011) 3017–3032, [1106.3366](#).
- [92] A. J. Ross, L. Samushia, C. Howlett, W. J. Percival, A. Burden, and M. Manera, “The clustering of the SDSS DR7 main Galaxy sample - I. A 4 per cent distance measure at $z = 0.15$,” *Mon. Not. Roy. Astron. Soc.* **449** (2015), no. 1 835–847, [1409.3242](#).
- [93] **BOSS** Collaboration, L. Anderson *et. al.*, “The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Baryon Acoustic Oscillations in the Data Release 10 and 11 galaxy samples,” *Mon.Not.Roy.Astron.Soc.* **441** (2014), no. 1 24–62, [1312.4877](#).
- [94] G. Efstathiou, “ H_0 Revisited,” *Mon.Not.Roy.Astron.Soc.* **440** (2014), no. 2 1138–1152, [1311.3461](#).
- [95] A. G. Riess *et. al.*, “A 3% Solution: Determination of the Hubble Constant with the Hubble Space Telescope and Wide Field Camera 3,” *Astrophys.J.* **730** (2011) 119, [1103.2976](#).
- [96] A. G. Riess *et. al.*, “A 2.4% Determination of the Local Value of the Hubble Constant,” [1604.01424](#).
- [97] **Planck** Collaboration, P. Ade *et. al.*, “Planck 2013 results. XVI. Cosmological parameters,” *Astron.Astrophys.* **571** (2014) A16, [1303.5076](#).
- [98] A. Barreira and Y. Dirian *in preparation*.
- [99] R. Trotta, “Bayes in the sky: Bayesian inference and model selection in cosmology,” *Contemp. Phys.* **49** (2008) 71–104, [0803.4089](#).
- [100] **BICEP2, Planck** Collaboration, P. Ade *et. al.*, “Joint Analysis of BICEP2/Keck Array and Planck Data,” *Phys. Rev. Lett.* **114** (2015) 101301, [1502.00612](#).
- [101] R. Trotta, “Applications of Bayesian model selection to cosmological parameters,” *Mon. Not. Roy. Astron. Soc.* **378** (2007) 72–82, [astro-ph/0504022](#).
- [102] S. Nesseris and J. Garcia-Bellido, “Is the Jeffreys’ scale a reliable tool for Bayesian model comparison in cosmology?,” *JCAP* **1308** (2013) 036, [1210.7652](#).
- [103] E.-J. Wagenmakers, T. Lodewyckx, H. Kuriyal, and G. Raoul, “Bayesian hypothesis testing for psychologists: a tutorial on the Savage-Dickey method,” *Cogn Psychol.* **60(3)** (2010) 158?189.
- [104] P. Mukherjee, J. Urrestilla, M. Kunz, A. R. Liddle, N. Bevis, and M. Hindmarsh, “Detecting and distinguishing topological defects in future data from the CMBPol satellite,” *Phys. Rev.* **D83** (2011) 043003, [1010.5662](#).

- [105] A. Oka, S. Saito, T. Nishimichi, A. Taruya, and K. Yamamoto, “Simultaneous constraints on the growth of structure and cosmic expansion from the multipole power spectra of the SDSS DR7 LRG sample,” *Mon.Not.Roy.Astron.Soc.* **439** (2014) 2515–2530, [1310.2820](#).
- [106] F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, G. B. Poole, L. Campbell, Q. Parker, W. Saunders, and F. Watson, “The 6dF Galaxy Survey: $z \approx 0$ measurements of the growth rate and σ_8 ,” *Mon.Not.Roy.Astron.Soc.* **423** (July, 2012) 3430–3444, [1204.4725](#).
- [107] C. Howlett, A. Ross, L. Samushia, W. Percival, and M. Manera, “The Clustering of the SDSS Main Galaxy Sample II: Mock galaxy catalogues and a measurement of the growth of structure from Redshift Space Distortions at $z = 0.15$,” [1409.3238](#).
- [108] C.-H. Chuang, F. Prada, F. Beutler, D. J. Eisenstein, S. Escoffier, *et. al.*, “The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: single-probe measurements from CMASS and LOWZ anisotropic galaxy clustering,” [1312.4889](#).
- [109] L. Samushia, B. A. Reid, M. White, W. J. Percival, A. J. Cuesta, *et. al.*, “The Clustering of Galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey (BOSS): measuring growth rate and geometry with anisotropic clustering,” *Mon.Not.Roy.Astron.Soc.* **439** (2014) 3504–3519, [1312.4899](#).
- [110] C. Blake, S. Brough, M. Colless, C. Contreras, W. Couch, *et. al.*, “The WiggleZ Dark Energy Survey: Joint measurements of the expansion and growth history at $z < 1$,” *Mon.Not.Roy.Astron.Soc.* **425** (2012) 405–414, [1204.3674](#).
- [111] S. de la Torre, L. Guzzo, J. Peacock, E. Branchini, A. Iovino, *et. al.*, “The VIMOS Public Extragalactic Redshift Survey (VIPERS). Galaxy clustering and redshift-space distortions at $z = 0.8$ in the first data release,” *Astron.Astrophys.* **557** (2013) A54, [1303.2622](#).
- [112] S. F. Daniel *et. al.*, “Testing General Relativity with Current Cosmological Data,” *Phys.Rev.* **D81** (2010) 123508, [1002.1962](#).
- [113] L. Amendola, M. Kunz, and D. Sapone, “Measuring the dark side (with weak lensing),” *JCAP* **0804** (2008) 013, [0704.2421](#).
- [114] **Euclid Theory Working Group** Collaboration, L. Amendola *et. al.*, “Cosmology and fundamental physics with the Euclid satellite,” *Living Rev.Rel.* **16** (2013) 6, [1206.1225](#).
- [115] R. Durrer and T. Kahniashvili, “CMB anisotropies caused by gravitational waves: A Parameter study,” *Helv.Phys.Acta* **71** (1998) 445–457, [astro-ph/9702226](#).
- [116] S. Weinberg, “Damping of tensor modes in cosmology,” *Phys.Rev.* **D69** (2004) 023503, [astro-ph/0306304](#).
- [117] R. Durrer, *The Cosmic Microwave Background*. Cambridge Univ. Press, Cambridge, 2008.

- [118] J. G. Williams, S. G. Turyshev, and D. H. Boggs, “Progress in lunar laser ranging tests of relativistic gravity,” *Phys. Rev. Lett.* **93** (2004) 261101, [gr-qc/0411113](#).
- [119] E. Babichev, C. Deffayet, and G. Esposito-Farese, “Constraints on Shift-Symmetric Scalar-Tensor Theories with a Vainshtein Mechanism from Bounds on the Time Variation of G ,” *Phys. Rev. Lett.* **107** (2011) 251102, [1107.1569](#).